

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2003, 10:29:00 ; Search time 4907 Seconds
(without alignments)
12413.129 Million cell updates/sec

Title: US-09-890-475-2

Perfect score: 3761

Sequence: 1 agtactacagtcacact.....aagtcgcggagagctctcg 3761

Scoring table:

IDENTITY_NUC
Gap 10.0 ; Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	614.8	16.3	635	17	B77833 T29E23TR TA
2	550.4	14.6	552	10	AV541057 AV541057
3	456.2	12.1	487	17	CNS00T03 AL090153 Arabidops
4	420.2	11.2	750	17	BH470959 BOHLJ19TR
5	335.2	8.9	781	17	BH479860 BOGRJ75TF
6	326.4	8.7	704	17	BH491272 BOHRH16TR

c	7	313.2	8.3	332	17	AL758973	AL758973 Arabidops
c	8	266.2	7.1	422	17	BH865818	BH865818 SALR_0999
c	9	230	6.1	250	17	AL758974	AL758974 Arabidops
c	10	211	5.6	408	9	AT2227805	AT2227805 Arabidops
c	11	200.8	5.3	439	17	BH753543	BH753543 SALR_0292
c	12	192.6	5.1	423	17	BH211901	BH211901 SALR_0068
c	13	191.6	5.1	251	17	BH753041	BH753041 SALR_0196
c	14	174.4	4.6	580	17	BH746972	BH746972 SALR_0068
c	15	165.4	4.4	471	9	AT992480	AT992480 701557920
c	16	161.8	4.3	785	14	BQ115269	BQ115269 EST600845
c	17	161.6	4.3	641	9	AA041118	AA041118 24384 CD4
c	18	134.6	3.6	557	12	BF050953	BF050953 EST436111
c	19	128.2	3.4	500	14	BO511650	BO511650 EST619065
c	20	128.2	3.4	671	12	BG890481	BG890481 EST516332
c	21	119.8	3.2	387	10	BE523453	BE523453 M7B4STM
c	22	119.8	3.2	685	10	AW738032	AW738032 EST339459
c	23	116.2	3.1	771	13	B1176101	B1176101 EST517184
c	24	114.8	3.1	428	17	BH633457	BH633457 SALR_0426
c	25	114.8	3.1	430	12	BG097426	BG097426 EST462041
c	26	109.6	2.9	637	14	BQ118418	BQ118418 EST603994
c	27	100.2	2.7	544	12	BF053940	BF053940 EST439170
c	28	87.6	2.3	474	12	BG044838	BG044838 saa33b07.
c	29	81.6	2.2	709	17	BH479870	BH479870 BOGRJ75TR
c	30	80.6	2.1	651	10	AW738612	AW738612 EST340039
c	31	79.6	2.1	430	12	BG097426	BG097426 EST462041
c	32	78.6	2.1	540	10	BE471538	BE471538 EST416391
c	33	72	1.9	1101	17	CNS00EVL	AL063706 Drosoph11
c	34	69	1.8	474	10	AW096504	AW096504 EST289684
c	35	66.6	1.8	511	10	BE611371	BE611371 sq80e12.y
c	36	66.6	1.8	996	17	CNS00FUR	AL071063 Drosoph11
c	37	66.4	1.8	1092	17	CNS020K7	AL175966 Tetrarodon
c	38	66	1.8	1225	17	CNS0161D	AL106171 Drosoph11
c	39	65.4	1.7	876	17	CNS009G1	AL033529 Drosoph11
c	40	64.2	1.7	905	17	CNS00KRX	AL077798 Drosoph11
c	41	63.2	1.7	811	17	CNS0083V	AL051088 Drosoph11
c	42	62.4	1.7	928	17	CNS00DKY	AL071865 Drosoph11
c	43	62.2	1.7	987	17	CNS014PQ	AL104456 Drosoph11
c	44	61.8	1.6	1200	17	CNS016CQ	AL106578 Drosoph11
c	45	61.6	1.6	1101	17	CNS00EQL	AL065526 Drosoph11

ALIGNMENTS

RESULT 1
LOCUS B77833/c
DEFINITION T29E23TR T29E23TR Arabidopsis thaliana genomic clone T29E23, DNA
ACCESSION B77833
VERSION B77833.1
KEYWORDS GI:2774472
SOURCE GSS.
ORGANISM thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 635)
Rounsley,S.D., Field,C.E., Baas,S., Linher,K., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter
J.C. End Sequence Database for Identifying Minimal Overlaps In
Arabidopsis genomic Sequencing. Update 3
Unpublished (1997)
Other GSSs: T29E23TR
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13 Reverse

Class: BAC ends
High quality sequence stop: 635.
Location/Qualifiers
1. .635
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T29E23"
/clone_1lb="TAMU"
/sex="hermaphrodite"
/note="Vector: BelobACII; Site_1: HindIII; Site_2: HindIII
; Produced by Rod Wang"

BASE COUNT 209 a 110 c 143 g 173 t
ORIGIN

Query Match 16.3%; Score 614.8; DB 17; Length 635;
Best Local Similarity 99.5%; Pred. No. 2.2e-130;
Matches 627; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 2587 CCCCCGAATATATGGTTCACCTCCACATGGTGGGTTAGGAAGAAGTATATGCAATG 2646
DB CCCCCGAATATATGGTTCACCTCCACATGGTGGGTTAGGAAGAAGTATATGCAATG 572
OY 2647 AACATCTGGCCCCAAATCATCTCTCCAGTCCAGGACATAGACTCATGACAGTACT 2706
DB 571 AACATCTGGCCCCAAATCATCTCTCCAGTCCAGGACATAGACTCATGACAGTACT 512
OY 2707 CTCGGCTTTGGTTACGAGACAGACATCCACTAGTACTCTCCCAATTCATGAGAC 2766
DB 511 CTCGGCTTTGGTTACGAGACAGACATCCACTAGTACTCTCCCAATTCATGAGAC 452
OY 2767 AACACAGTTACCATATGATATACAAAGGTTTACAGACATTCACCATTGGAAGAAGAT 2826
DB 451 AACACAGTTACCATATGATATACAAAGGTTTACAGACATTCACCATTGGAAGAAGAT 392
OY 2827 AATTGGGTTTATCCAAATCAAGGCTCTCCGAGTAATCATATTAAGCCCAATAGG 2886
DB 391 AATTGGGTTTATCCAAATCAAGGCTCTCCGAGTAATCATATTAAGCCCAATAGG 332
OY 2887 AGGAATGTAATTTGTAACAAAGCTTTTGTTCCTTAAGTATGCTATTTTAACT 2946
DB 331 AGGAATGTAATTTGTAACAAAGCTTTTGTTCCTTAAGTATGCTATTTTAACT 272
OY 2947 CCCAAGCTCTCAAAATTAATTAATGTTGGGGTTAAGAATGCAAAATTTTTTGGTC 3006
DB 271 CCCAAGCTCTCAAAATTAATTAATGTTGGGGTTAAGAATGCAAAATTTTTTGGTC 212
OY 3007 CTGTAAATGACATTAAGATGCTAATGTAATGCTTCAGAGCTTTTAACTCAACCTCAAT 3066
DB 211 CTGTAAATGACATTAAGATGCTAATGTAATGCTTCAGAGCTTTTAACTCAACCTCAAT 152
OY 3067 ACATGATATGACATTAATGACCTCTGGGCTTGGTCATCTGAT -TCTCTTCATC 3125
DB 151 ACATGATATGACATTAATGACCTCTGGGCTTGGTCATCTGATCTGCTTCATC 92
OY 3126 TTCTGCTCTGTTCTCTTCTTCTGCTGCACTGCTCGAGCAATTCGGAATTCACACT 3185
DB 91 TTCTGCTCTGTTCTCTTCTTCTGCTGCACTGCTCGAGCAATTCGGAATTCACACT 32
OY 3186 TGTGCTTACAGTTTCCATGACAGAGCTT 3215
DB 31 TGTGCTTACAGTTTCCATGACAGAGCTT 2

RESULT 2
AV541057/c 552 bp mRNA linear EST 07-SEP-2000
LOCUS AV541057 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
DEFINITION cDNA clone RZ159d09F 3', mRNA sequence.
ACCESSION AV541057
VERSION AV541057.1 GI:8702815
KEYWORDS EST.
SOURCE thale cress.

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 552)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp. URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1. .552
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZ159d09F"
/clone_1lb="Arabidopsis thaliana roots Columbia"
/issue_type="roots"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 174 a 97 c 117 g 164 t
ORIGIN

Query Match 14.6%; Score 550.4; DB 10; Length 552;
Best Local Similarity 99.8%; Pred. No. 1.3e-115;
Matches 551; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2516 AGATGAATATACAGCTCTTGTGAGTATGTAATGCTGAGCCGTCACATCTTTCCATCG 2575
DB 552 AGATGAATATACAGCTCTTGTGAGTATGTAATGCTGAGCCGTCACATCTTTCCATCG 493
OY 2576 CTCAAGAAATCCCCGAATATATGTTCCACTCCACATGGTGGTTAGGAAGAAGT 2635
DB 492 CTCAAGAAATCCCCGAATATATGTTCCACTCCACATGGTGGTTAGGAAGAAGT 433
OY 2636 ATATGATATGACATCTGCGCCCAATTCATCACTCCAGTCCAGGACATAGACTTCA 2695
DB 432 ATATGATATGACATCTGCGCCCAATTCATCACTCCAGGACATAGACTTCA 373
OY 2696 TCGACATGACTCTCCGCTTGTGTCACGACAGACATCCACATCACTACTCTCTCC 2755
DB 372 TCGACATGACTCTCCGCTTGTGTCACGACAGACATCCACATCACTACTCTCTCC 313
OY 2756 AATTCATGACACACACAGTTACCATATGTTATACAAAGGTTTACAGCATTCACCATC 2815
DB 312 AATTCATGACACACACAGTTACCATATGTTATACAAAGGTTTACAGCATTCACCATC 253
OY 2816 TGAAGAAGATATTTGGGTTTATCAATCAAGGTCCTCGAGTAAGTCAATCATTTAGA 2875
DB 252 TGAAGAAGATATTTGGGTTTATCAATCAAGGTCCTCGAGTAAGTCAATCATTTAGA 193
OY 2876 CCCCCAATAGAGAGATATTAATTTGTACAAAAGCTTTTGTGTTTGTCTTAAGTATGTA 2935
DB 192 CCCCCAATAGAGAGATATTAATTTGTACAAAAGCTTTTGTGTTTGTCTTAAGTATGTA 133
OY 2936 TTTTATTTAATCCCAAGTCTCAAAATTTAATTAATGTTGGGCTTAAGAATGCAAA 2995
DB 132 TTTTATTTAATCCCAAGTCTCAAAATTTAATTAATGTTGGGCTTAAGAATGCAAA 73
OY 2996 TTTTGTGCTCTGTAATGACATTTAAGATGCTAATGTAATGCTTCAGAGGCTTTAGT 3055
DB 72 TTTTGTGCTCTGTAATGACATTTAAGATGCTAATGTAATGCTTCAGAGGCTTTAGT 13
OY 3056 CAACCTCAGATA 3067
DB 12 CAACCTCAGATA 1

RESULT 3	CNS00TU3/c	487 bp	DNA	Linear	GSS 28-JUN-1999
LOCUS	CNS00TU3				
DEFINITION	Arabidopsis thaliana genome survey sequence SP6 end of BAC T572 of TBMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.				
ACCESSION	AL090153				
VERSION	AL090153.1	GI:5291293			
KEYWORDS	GSS.				
SOURCE	Arabidopsis thaliana.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Rosidae; eucotsids II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	1 (bases 1 to 487) Salanoubat,M., Choinsne,N., Attigunave,F., Brotlier,P., Wincker,P., Samsou,D., Saurin,W., Weissenbach,J. and Quetier,F.				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 487)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequelegenoscope.cns.fr)				
FEATURES	- Web : www.genoscope.cns.fr				
SOURCE	Location/Qualifiers				
	1..487				
	/organism="Arabidopsis thaliana"				
	/strain="Columbia"				
	/db_xref="taxon:3702"				
	/clone="T572"				
	/clone_lib="TBMU"				
	/note="end : SP6"				
BASE COUNT	164 a 114 c 77 g 132 t				
ORIGIN					
Query Match	12.1% ; Score 456.2; DB 17; Length 487;				
Best Local Similarity	96.2% ; Pred. No. 5,4e-94;				
Matches 484; Conservative	0; Mismatches 3; Indels 16; Gaps 1;				
OY	1360 GCTGGGGCGGAGAAAATGGATGCCAAGGGGTTGCTTACTAGTGGCTTTGGTGTGTT	1419			
DB	487 GCTGGGGCGGAGAAAATGGATGCCAAGGGGTTGCTTACTAGTGGCTTTGGTGTGTT	428			
OY	1420 CTTTCAAACTTTAGAGAGTACAGATTTCGCGATTGATTAAGATGATGGTTGCAATG	1475			
DB	427 CTTTCAAACTTTAGAGAGTACAGATTTCGCGATTGATTAAGATGATGGTTGCAATG	368			
OY	1480 AATGCGCGGTGCTTTGAAGCGGTACACAGTTCTTGTCCTATGAGTCTACAGTCCATATTC	1533			
DB	367 AATGCGCGGTGCTTTGAAGCGGTACACAGTTCTTGTCCTATGAGTCTACAGTCCATATTC	324			
OY	1540 TGTTCACATCGGTGATTTCAATTCGAAGGGTTCCTTTGTTGACATCATCGACCA	1599			
DB	323 TGTTCACATCGGTGATTTCAATTCGAAGGGTTCCTTTGTTGACATCATCGACCA	264			
OY	1600 CATCAAGTTCCATCTTTGTTTTCGATTAAGCTTTGATGTATAAATCTAGAGAGACATCA	1655			
DB	263 CATCAAGTTCCATCTTTGTTTTCGATTAAGCTTTGATGTATAAATCTAGAGAGACATCA	204			
OY	1660 AATATTAGAGTCAATGACTGATTTAGGCCAAATCTAGCTAGAAATTAATCTGGAAAGA	1711			
DB	203 AATATTAGAGTCAATGACTGATTTAGGCCAAATCTAGCTAGAAATTAATCTGGAAAGA	144			
OY	1720 ACTTGGACATCCTCAACCATAGTTTGGTACGAAATTTGTTGCTTGCACACCAATGAT	1777			
DB	143 ACTTGGACATCCTCAACCATAGTTTGGTACGAAATTTGTTGCTTGCACACCAATGAT	84			
OY	1780 AGGCTATTGCTTGAATAGTGTCTTGTGGTTTCCAAATATTGGAAGTTAAATGTAT	1833			
DB	83 AGGCTATTGCTTGAATAGTGTCTTGTGGTTTCCAAATATTGGAAGTTAAATGTAT	24			
OY	1840 GACTTAGCTGTGGATTAATTT 1862				

[illegible]

Db	328	CTCACAACATTAGAACCCACAGAGAAAGACAGTACAGATGGCAGATATCAAAAGACAAATGATTA	269
Qy	2302	GCCTGGAGAAAGACACTCTTCAAGCTCCGACAAAGAGATGAGAAAGAAAGCAATCTCTCA	2361
Db	268	AGTTGAGAGAAACATCTTTCTTAACCTCGACAAACAGATGAGAGGAGGAAAGATCCATCA	209
Qy	2362	GTTTAATGAGAGAACCCGACTTGGCCAGAGAAATGTAAACCAAGATAAACGTCGA	2421
Db	208	GTTTAATGAGAGAAACAGATGAGAGAGATGATTAACCAAGATGAAACGTCGA	149
Qy	2422	GCTGTGACCCATGGAATATGACACAGATTAATCTTATGCTATCTCTACCTACCCG	2481
Db	148	GGTTGTCACATGAGAAATGCCACACGAGCTTCATCTATCTCTACCTACCCG	89
Qy	2482	ATAGAACTTTCTTAGTCAAGAGACGATGACCAAGATGAATGAATATGAGCTTGTGAGTA	2541
Db	88	CCGAGAGCTTCCTGATGTCACAGAGAC-----GATGAATATATGAGCTTGTGAGTA	38
Qy	2542	GTTTACCTGGCCGTCACACATCTTTTCTCACTGCTC	2578
Db	37	GTTTACCTGAGCCTTACACAGGTTTCTCATGCGTC	1
RESULT 5			
LOCUS	BH479860	781 bp	DNA
DEFINITION	BOGRJ75FR BOGR Brassica oleracea genomic clone BOGRJ75, DNA		
ACCESSION	BH479860		
VERSION	BH479860.1	GI:17687964	
KEYWORDS	GSS.		
SOURCE	Brassica oleracea.		
ORGANISM	Brassica oleracea		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.		
AUTHORS	town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.		
TITLE	Whole genome shotgun sequencing of Brassica oleracea		
JOURNAL	Unpublished (2001)		
COMMENT	Other GSSs: BOGRJ75FR		
	Contact: Chris Town		
	TIGR		
	9112 Medical Center Drive, Rockville, MD 20850, USA.		
	Tel: 301-838-3523		
	Fax: 301-838-0208		
	Email: cdtown@tigr.org		
	DNA is from a doubled haploid provided by Tom Osborn.		
	Seq primer: TF		
	Class: sheared ends.		
FEATURES			
source	location/Qualifiers		
	1..781		
	/organism="Brassica oleracea"		
	/strain="T01000DH3"		
	/db_xref="taxon:3712"		
	/clone="BOGRJ75"		
	/clone_1lb="BOGR"		
	/note="Vector: PHOS1, site_1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"		
BASE COUNT	188 a 179 c 239 g 175 t		
ORIGIN			
Query Match	8.9%; Score 335.2; DB 17; Length 781;		
Best Local Similarity	67.8%; Pred. No. 3.3e-66;		
Matches 519; Conservative	0; Mismatches 208; Indels 39; Gaps 2;		
Qy	655	CGACGAGAAAGATTTACCGAATGTTCTTCAAAACAGAGCTCAAGATATGACATATACGATC	714
Db	19	CGACGAGAGAGAGCGAGGCTACTGTCCAAAACCGCTCCAAACATCAGAAACACAGATC	78
Qy	715	GGTCATCTTACGACGCTCAATTTTGAATTCATAGACGATTTAGCTGCTTTTACGTT	774
Db	79	GAAACATCTTACGACGCTCAATTTTGAATTCATAGACGATTTAGCTGCTTTTACGTT	138

[illegible]

FEATURES		source	location/Qualifiers
		1..704	/organism="Brassica oleracea"
			/strain="T0100DBH3"
			/db_xref="taxon:3712"
			/clone="BOHBH16"
			/clone_11b="BOHB"
			/note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT		173 a	165 c 157 g 209 t
ORIGIN			
Query Match	8.7%; Score 326.4; DB 17; Length 704;		
Best Local Similarity	72.8%; Pred. No. 3.4e-64;		
Matches 527; Conservative	0; Mismatches 111; Indels 56; Gaps 6;		
Db	2161 TTGATATATTTATTTCTTCACACGAAAGAGCG-6CTACAAAGCAGCTAGCTGTATATCA	2219	
Qy	TTGATATATTTATTTCTTCACACGAAAGAGCG-6CTACAAAGCAGCTAGCTGTATATCA	2219	
Db	695 TTGGCTTTTATTTTCTTCACACGAAAGAGCTGGAAGCAAGAGTGGATGCTATCA	636	
Qy	2220 TCAGTTATGACGTGTATGAGACCTCACAGTTAGATCTCTGGAAGAAATACCAAGATG	2279	
Db	635 TCAGTTATGACGTGTATGAGACCTCACAGTTAGATACCAAGAAAGAAATACCAAGATG	576	
Qy	2280 CAGATCAAGAGCAAAATTTGTTAGCTTGAGAAAGACACTCTTCAGCTCGACAAAGAGAT	2339	
Db	575 CAGATCAAGAGCAAAATTTGTTAGCTTGAGAAAGAGATTTGTTAGCTCGACAAAGAGAT	516	
Qy	2340 GAA-----GAGAAACAGATCTCCAGTTTAAATGAGAA	2375	
Db	515 GAAAGAGCGAGATCCATCATGTCGAATGAGAGAGCGGATCATCATAGCTTAAGGAGAA	456	
Qy	2376 GCGGCACTTGGCCACAGAGATGTATATACCAAGATATAAAAGCTCCAGGTTGTCCACAT	2435	
Db	455 GCGGCAATTAAGAGAGATGTATATACCAAGATATAAAAGCTCCAGGTTGTCCAGAAAG	396	
Qy	2436 GAAATGCCACAGTACTTCTTCATCGTATTTCTCTATCTACCTGATAGAGCTTCTCT	2495	
Db	395 GAAATGCCACAGTACTTCTTCATCGTATTTCTCTATCTACCTGATAGAGCTTCTCTCT	336	
Qy	2496 AGTCAAGAGCAGATGACCAAGATGAATATATCAGCTCTTGAGAGTATACCTCGGCGCG	2555	
Db	335 AGTCAAGAGCAGATGACCAAGATGAATATATCAGCTCTTGAGAGTATACCTCGGCGCG	276	
Qy	2556 TCAACATCTTTCTCTATCGCTCAAGAAAGATCCCGGAATATATGCTTCCACTTCCACAT	2615	
Db	275 TCAACATCTTTCTCTATCGCTCAAGAAAGATCCCGGAATATATGCTTCCACTTCCACAT	216	
Qy	2616 GGTGGTTAGGAAGAAAGCTATATGCAATATGAACATCTGGGCCCAATTTACTACTCTCA	2675	
Db	215 GGTGGTTAGGAAGAAAGCTCTCTCGATGATGATCAACCTCCCAATTTCTTATTTCTCG	156	
Qy	2676 GGTCAAGGACATAGACTCTCATGAGACAGTACTCTCGCTTTGGTTCACGAGAGAGACAT	2735	
Db	155 G-----TTTCAAGAAAGTACTCTCC-----AGTTCACGAGACAGAGACTT	117	
Qy	2736 CCAGTCACTACTCTCTCTCCATTCATATGACAAACAAGATTAAGCATATGATATCAAGG	2795	
Db	116 CCCTCAAGATTAATCTCTCTCCATTCATATGAGGACAAACAAGATTAAGCATATGATATCA	63	
Qy	2796 GTTTACAGACATTCACCATCTGAAGAAATATTTGGTTATTCACATCAAGAGCTCTCT	2855	
Db	62 ---TAAAGCATTCACCATCTG-TGAAGATATCTTGGCTTTTCCAAATCAGAGACTCT	7	
Qy	2856 CGCA 2859		
Db	6 CGTA 3		

ACCESSION	AL758973	genomic survey sequence.
VERSION	AL758973.1	GI:21497321
KEYWORDS	GSS.	
SOURCE	thale cress.	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
TITLE	A pipeline for automated high-throughput generation of FSNs	
JOURNAL	(flanking sequence tags) from Arabidopsis thaliana T-DNA	
AUTHORS	transformed lines	
REFERENCE	Unpublished	
TITLE	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
JOURNAL	1 Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saadler,H.	
AUTHORS	2 and Weisshaar,B.	
REFERENCE	3 A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)	
TITLE	for flanking sequence tag based reverse genetics	
JOURNAL	Unpublished	
AUTHORS	3 (bases 1 to 332)	
REFERENCE	Strizhov,N., Rosso,M., Li,Y. and Weisshaar,B.	
TITLE	Direct Submission	
JOURNAL	Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer	
AUTHORS	Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany	
REFERENCE	This sequence is recovered from the left border of the T-DNA. It	
TITLE	indicates an insertion close to or within gene At4g00650. The	
JOURNAL	sequences are generated at the MPI for Plant Breeding Research in	
AUTHORS	the context of the GABI-Kat project. GABI-Kat is part of the German	
REFERENCE	Plant Genomics program designated 'GABI'. Information on line	
TITLE	availability can be found at:	
JOURNAL	http://www.mpiz-koeln.mpg.de/GABI-Kat/.	
AUTHORS	Location/Qualifiers	
REFERENCE	1..332	
TITLE	/organism="Arabidopsis thaliana"	
JOURNAL	/strain="Columbia 0"	
AUTHORS	/db.xref="taxon:3702"	
REFERENCE	/clone="GK-16AE05-013270"	
TITLE	/clone_lib="Arabidopsis thaliana T-DNA insertion lines"	
JOURNAL	/note="PCR was performed on DNA from Arabidopsis thaliana	
AUTHORS	plants (T1) which were transformed with the T-DNA from	
REFERENCE	vector pAC161. The lines contain one or more T-DNA	
TITLE	insertions. The DNA fragment(s) resulting from the PCR	
JOURNAL	were directly sequenced to determine the genomic sequence	
AUTHORS	flanking the insertion. Sequences displaying significant	
REFERENCE	similarity to the A. thaliana nuclear genome sequence were	
TITLE	processed for submission. T-DNA derived sequences were	
JOURNAL	removed"	
AUTHORS	BASE COUNT 108 a 70 c 50 g 103 t 1 others	
REFERENCE	ORIGIN	
TITLE	Query Match 8.3%; Score 313.2; DB 17; Length 332;	
JOURNAL	Best Local Similarity 99.1%; Pred. No. 3,4e-61;	
AUTHORS	Matches 315; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
REFERENCE	1694 CCTAGCTAGAATTTATCTGTGGAAGAACTTGCAACTCTCAACCATAGGTTTGGTAGCA 17533	
TITLE	32A CATTACTGGAATTAATTCATCGGAAAACAATTGGAACCTCAACCAAGATTGGTAGCAA 265	
JOURNAL	1754 ATTGTGCTGTGCAGAACCAATGATAGGCTATTGCCCTTGAAATAGTGTCTGTGGCTT 18133	
AUTHORS	264 ATTTGTGCTGTGCAGAACCAATGATAGGCTATTGCCCTTGAAATAGTGTCTGTGGCTT 205	
REFERENCE	1814 TCCAATATTGGGAAGTTAAATCGTATGACTTAGCTGTGGATACATAATTAAGCTTAACA 18737	
TITLE	204 TCCAATATTGGGAAGTTAAATCGTATGACTTAGCTGTGGATACATAATTAAGCTTAACA 145	
JOURNAL	1874 ATGCCAAGCTAAGAAAGTGTACTTACACATATTTCTATGGTCAATAGCTATAGCTTAAT 19333	
AUTHORS	144 ATGCCAAGCTAAGAAAGTGTACTTACACAAATATTTCTATGGTCAATAGCTATAGCTTAAT 85	
REFERENCE	1934 CAAGTATCAAGGCTGAATGCATATTATTAAGGCTTGAAGATGATTATACCTTTGGCANG 19933	

QY 1765 TCAGAACCAATGATGCTATTCCTGGAATAGCTTCCTGGCTTCGCAATATGCG 1824
 |||||
 DB 118 TCAGAACCAATGATGCTATTCCTGGAATAGCTTCCTGGCTTCGCAATATGCG 59
 |||||
 QY 1825 AAGTTAAATCGTACTAGCTGCTGGATGATTAATTAAGCTTAAGCAATGCCACT 1882
 |||||
 DB 58 AAGTTAAATCGTACTAGCTGCTGGATGATTAATTAAGCTTAAGCAATGCCACT 1

RESULT 10
 AU227805 408 bp mRNA linear EST 23-APR-2002
 LOCUS AU227805
 DEFINITION AU227805 RAFL15 Arabidopsis thaliana CDNA clone RAFL15-24-J05 3',
 mRNA sequence.
 ACCESSION AU227805
 VERSION AU227805.1 GI:19742452
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 408)
 AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
 Akiyama,K., Enju,A., Oono,Y., Sekurai,T., Carninci,P., Kawai,J.,
 Itoh,M., Ishii,Y., Arai,K., Shibata,K., Shinagawa,A., Muramatsu,
 M., Hayashizaki,Y. and Shinozaki,K.
 Large scale analysis of Arabidopsis full-length cDNA
 Unpublished (2002)
 CONTACT: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified pluescript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/Plant/index_e.html) for further
 details.

FEATURES
 source location/Qualifiers
 1..408
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="RAFL15-24-J05"
 /clone_lib="RAFL15"
 /issue_type="mixture of silique and flower"
 /lab_host="DH10B"
 /note="Site_1: BamHI; Site_2: SalI"
 /lab_host="DH10B"
 /note="Site_1: BamHI; Site_2: SalI"

BASE COUNT 101 a 87 c 77 g 143 t
 ORIGIN

Query Match 5.6%; Score 211; DB 9; Length 408;
 Best Local Similarity 91.8%; Pred. NO. 1e-37;
 Matches 223; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 2993 AAATTTTTCCTGCTTAATGACATTAAGATGCTTAATGCTTCACAGGTTT 3052
 |||||
 DB 4 AAAAATTTTTCCTGCTTAATGACATTAAGATGCTTAATGCTTCACAGGTTT 63
 |||||
 QY 3053 AGTCAACCTGAGATACATGATACATTAATTAATGACCTGCTGCTGCTCATCTG 3112
 |||||
 DB 64 AGTCAACCTGAGATACATGATACATTAATTAATGACCTGCTGCTGCTCATCTG 123
 |||||
 QY 3113 GATCTCTATCT 3172
 |||||
 DB 124 GATCTCTATCT 183
 |||||
 QY 3173 CGAATTCACACCTGCTGCTTACATTTCCATGACACAGCTTTCCATGAATGATTTA 3232

DB 184 GGAATTCACACCTGCTTACAGTTCCATGACACAGCTTTCCATGAATGATTTA 243
 |||||
 QY 3233 TGT 3235
 |||||
 DB 244 TGT 246

RESULT 11
 BH753543 439 bp DNA linear GSS 27-FEB-2002
 LOCUS BH753543
 DEFINITION SALK_029287.54.50.x Arabidopsis thaliana TDNA insertion lines
 Arabidopsis thaliana genomic clone SALK_029287.54.50.x, DNA
 sequence.
 ACCESSION BH753543
 VERSION BH753543.1 GI:18973692
 KEYWORDS GSS.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 439)
 AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadriab,
 C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednals,L., Shinn,P.,
 Zimmerman,J. and Ecker,J.R.
 A Sequence-indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 Unpublished (2001)
 CONTACT: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGNAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: eckers@salk.edu
 This is single pass sequence recovered from the left border of
 TDNA. This sequence lies within 300 bases of the 5' end of
 At5g51090 and 300 bases of the 3' end of At5g51100.
 Class: TDNA tagged.

FEATURES
 source location/Qualifiers
 1..439
 /organism="Arabidopsis thaliana"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="SALK_029287.54.50.x"
 /clone_lib="Arabidopsis thaliana TDNA insertion lines"
 /note="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT 91 a 103 c 122 g 115 t 8 others
 ORIGIN

Query Match 5.3%; Score 200.8; DB 17; Length 439;
 Best Local Similarity 90.7%; Pred. NO. 2.2e-35;
 Matches 225; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 2968 TTTAATGTTTGGGCTTAAGATGCAATTTTTCGTC-CTGTAATGACATTAAGAT 3026
 |||||
 DB 71 TATAGTAGTGTATTAAGATGCAAAAAATTTTCGTCGTGTAATTAACATTAAGAT 130
 |||||
 QY 3027 GCTAATGTTATGCTTCAGAGTTTTCAGTCAACCTGACATGATGATGATGATTA 3086
 |||||
 DB 131 GCTAATGTTATGCTTCAGAGTTTTCAGTCAACCTGACATGATGATGATGATTA 190
 |||||
 QY 3087 ATAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3146
 |||||
 DB 191 ATAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 250
 |||||
 QY 3147 TTCTGTTGCACTGCTGAGCAATTCGGAATTCGAACCTGTGCTTAAGTTCCATGA 3206

Db 251 TTTCTTTGACACCTGCTGAGCAATGGGATTCACCACTTGTCTTACAGTTTCCACATA 310
 QY 3207 CACAAGCT 3214
 Db 311 CACAAGCT 318

RESULT 12

LOCUS BH211901 423 bp DNA linear GSS 24-OCT-2001
 DEFINITION SALK_006814 Arabidopsis thaliana TDNA insertion lines Arabidopsis
 thaliana genomic clone SALK_006814, DNA sequence.

ACCESSION BH211901
 VERSION BH211901.1 GI:16392653

KEYWORDS GSS.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 423)
 Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadgilab
 ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
 , Zimmermann,J. and Ecker,J.R.
 A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 Unpublished (2001)

TITLE Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: eckersalk.edu

COMMENT This is single pass sequence recovered from the left border of
 TDNA. This sequence lies within 300 bases of the 5' end of
 At5g51090 and 300 bases of the 3' end of At5g51100.
 Class: TDNA tagged.

FEATURES
 source Location/Qualifiers
 1..423
 /organism="Arabidopsis thaliana"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="SALK_006814"
 /note="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at http://signal.salk.edu/tdna_protocols.html."

BASE COUNT 72 a 145 c 73 g 123 t 10 others
 ORIGIN

Query Match 5.1%; Score 192.6; DB 17; Length 423;
 Best Local Similarity 95.7%; Pred. No. 1.7e-33;
 Matches 198; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3008 TGTATTCACATTTAAGATGCTAATGTTATTTGCTTCAGAGTTTGTCAACCTCGAGTA 3067
 Db 1 TGTATTCACATTTAAGATGCTAATGTTATTTGCTTCAGAGTTTGTCAACCTCGAGTA 60
 QY 3068 CATGATATTCACATTTAAGATGCTTCGCTCTGGTGCATGATGATCTCTTCATCTT 3127
 Db 61 CATGATATTCACATTTAAGATGCTTCGCTCTGGTGCATGATGATCTCTTCATCTT 120
 QY 3128 CTGCTCTGCT 3187
 Db 121 CTGCTCTGCT 180
 QY 3188 TGTTCACAGTTTCCATGACACAAAGCT 3214
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 181 TGCTTACAGTTTCCATGACACAAAGCT 207

RESULT 13

LOCUS BH753041 251 bp DNA linear GSS 27-FEB-2002
 DEFINITION SALK_019687.56.00.x Arabidopsis thaliana TDNA insertion lines
 Arabidopsis thaliana genomic clone SALK_019687.56.00.x, DNA
 sequence.

ACCESSION BH753041
 VERSION BH753041.1 GI:18972567

KEYWORDS GSS.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 251)
 Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadgilab
 ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
 , Zimmermann,J. and Ecker,J.R.
 A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 Unpublished (2001)

TITLE Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: eckersalk.edu

COMMENT This is single pass sequence recovered from the left border of
 TDNA. This sequence lies within 300 bases of the 5' end of
 At5g51090 and 300 bases of the 3' end of At5g51100.
 Class: TDNA tagged.

FEATURES
 source Location/Qualifiers
 1..251
 /organism="Arabidopsis thaliana"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="SALK_019687.56.00.x"
 /note="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at http://signal.salk.edu/tdna_protocols.html."

BASE COUNT 55 a 66 c 45 g 85 t
 ORIGIN

Query Match 5.1%; Score 191.6; DB 17; Length 251;
 Best Local Similarity 95.6%; Pred. No. 2.7e-33;
 Matches 197; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3009 GTAATTCACATTTAAGATGCTAATGTTATTTGCTTCAGAGTTTGTCAACCTCGATAC 3068
 Db 1 GTAATTCACATTTAAGATGCTAATGTTATTTGCTTCAGAGTTTGTCAACCTCGATAC 60
 QY 3069 ATGATATTCACATTTAAGATGCTTCGCTCTGGTGCATGATGATCTCTTCATCTTC 3128
 Db 61 ATGATATTCACATTTAAGATGCTTCGCTCTGGTGCATGATGATCTCTTCATCTTC 120
 QY 3129 TGTCTCTGCT 3188
 Db 121 TGTCTCTGCT 180
 QY 3189 GCTTACAGTTTCCATGACACAAAGCT 3214
 Db 181 GCTTACAGTTTCCATGACACAAAGCT 206
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 14

LOCUS	BH746972	580 bp	DNA	linear	GSS 27-FEB-2002
DEFINITION	SALK_006812.29.99.f Arabidopsis thaliana TDNA insertion lines				
ACCESSION	BH746972				
VERSION	BH746972				
KEYWORDS	Arabidopsis thaliana genomic clone SALK_006812.29.99.f, DNA sequence.				
SOURCE	BH746972.1	GI:18960087			
ORGANISM	GSS.				
REFERENCE	thale cress.				
AUTHORS	Arabidopsis thaliana Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 580) Alonso,J.M., Leisse,T.J., Batajes,P., Chen,H., Cheuk,R., Gadri nab, ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., ,Zimmerman,J., and Ecker,J.R. A Sequence-Indexed Library of Insertion Mutations In the Arabidopsis Genome Unpublished (2001)				
TITLE	Contact: Joseph R. Ecker				
JOURNAL	Salk Institute Genomic Analysis Laboratory (SIGNAL)				
COMMENT	The Salk Institute for Biological Studies 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA Tel.: 858 453 4100 x1752 Fax: 858 558 6379 Email: eckeresalk.edu This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of Atg5j1090 and 300 bases of the 3' end of At5g51100. Class: TDNA tagged.				
FEATURES	Location/Qualifiers				
SOURCE	1..580				
	/organism="Arabidopsis thaliana"				
	/strain="Columbia 0"				
	/db_xref="taxon:3702"				
	/clone="SALK_006812.29.99.f"				
	/clone_1lb="Arabidopsis thaliana TDNA insertion lines"				
	/note="PCR was performed on Arabidopsis thaliana lines				
	each of which contains one or more TDNA insertion				
	elements. The resultant fragment for each line was				
	directly sequenced to determine the genomic sequence at				
	the site of insertion. Details of the protocols used can				
	be found at http://signal.salk.edu/cdna_protocols.html "				
BASE COUNT	113 a 208 c 98 g 158 t 3 others				
ORIGIN					
Query Match	4.6%; Score 174.4; DB 17; Length 580;				
Best Local Similarity	92.0%; Pred. No. 2,6e-29;				
Matches 184; Conservative	0; Mismatches 16; Indels 0; Gaps 0;				
QY	3016 ACATTAAAGATGCTTAATGTATTGCTTCAGAGGTTTATGTCACCACTCGATACATCATGATA 3075				
DB					
	168 ACAAATGAAGAATGATATGATATGCTTATGCTTACCAGGTTTTTGTCACCGCGAATACATCATGATG 227				
QY	3076 TCACTATCAATTAATACCTGGCCCTTGTCATGCATATGATATGCTTCATCTCTGCTCTCT 3135				
DB					
	228 TCACTATCTAATTAATACCTCTGCTACTTCATCATATTCGATTCCTCTTCAATCTTCCTCTCT 287				
QY	3136 GTTCCTTCTTGTCTCGTGTGCACTGCTCGAGCAATTCGGGATTCACAACCTTGCTGTACA 3195				
DB					
	288 GTTCCTTCTTGTCTCTGCTTGCACCTGCTCGAGCAATTCGGGATTCACAACCTTGCTGTACA 347				
QY	3196 GTTCCCATGACACAAGCTT 3215				
DB					
	348 GTTCCCATGACACAAGCAT 367 .				
RESULT 15					
LOCUS	AI992480	471 bp	mRNA	linear	EST 08-SEP-1999
DEFINITION	701557920 A. thaliana, Ohio State clone set Arabidopsis thaliana				
	cdna clone 701557920, mRNA sequence.				

```

ACCESSION      AI992480
VERSION        AI992480.1
KEYWORDS       GI:5839385
SOURCE         thale cress.
ORGANISM       Arabidopsis thaliana
REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
AUTHORS        1 (bases 1 to 471)
               Chen,J., Montanya,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,
               Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,
               Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D., Fan,R.,
               Rose,M., Warren,B., Ton,B., Kastury,R., Borillo,C., Carpio,T.,
               Policky,J., Suzuki,G., Argentine,C., Shah,S., Nobrigha,A., Murry,L.,
               Turner,C., Kirkorian,S., Elder,L. and Hanson,D.
               Arabidopsis thaliana Gene Expression Microarray
               Unpublished (1999)
JOURNAL        Contact: David Smoller, Ph.D.
               Genome Systems, Inc., a wholly owned subsidiary of Incyte
               Pharmaceuticals, Inc.
               4633 World Parkway Circle, St. Louis, MO 63134, USA
               Tel: 877-577-2733
               Fax: 314-427-3324
               Email: service@genomesystems.com.
COMMENT        FEATURES
               source
               1..471
               /organism="Arabidopsis thaliana"
               /db_xref="taxon:3702"
               /clone="701557920"
               /clone.lib="A. thaliana, Ohio State clone set"
               /note="cDNA library was made from selected clones from the
               Arabidopsis thaliana Ohio State clone set."
BASE COUNT    117 a 118 c 83 g 153 t
ORIGIN
Query Match   4.4%; Score 165.4; DB 9; Length 471;
Best Local Similarity 96.3%; Pred. No. 3e-27;
Matches 180; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY            3050 TTTAGTCACCTCAG-ATACATGATATCACTAATCTAATAGACCTGTGGCTTTGGTCA 3108
               |||||
DB            1 TTTAGTCACCTCAGTATACATGATATCACTAATCTAATAGACCTGTGGTCAATCA 60
QY            3109 TCTGATTCCTTCATCTCTGTCGCTGTCCTTGTCTGTCGTCGTCGTCGACG 3168
               |||||
DB            61 TCTGATTCCTTCATCTCTGTCGCTGTCCTTGTCTTGTCTTGGCAGCTCGACGA 120
QY            3169 ATTGGGATTCACACCTGTGCTTACACTTCCATGACACAGCTTTTCATGAATGA 3228
               |||||
DB            121 ATTGGGATTCACACCTGTGCTTACACTTCCATGACACAGCTTTTCATGAATGA 180
QY            3229 TTTATGT 3235
               |||||
DB            181 TTTATGT 187

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2003, 10:29:00 ; Search time 9421 Seconds
(without alignments)

11618.263 Million cell updates/sec

Title: US-09-890-475-2

Perfect score: 3761
Sequence: 1 agactacacaagtcacaact.....aagltgcgcgagagctctcg 3761

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl :
1: gb_da : *
2: gb_hcg : *
3: gb_in : *
4: gb_cm : *
5: gb_ov : *
6: gb_pat : *
7: gb_ph : *
8: gb_pl : *
9: gb_pr : *
10: gb_ro : *
11: gb_sts : *
12: gb_sy : *
13: gb_un : *
14: gb_vl : *
15: em_da : *
16: em_fun : *
17: em_hum : *
18: em_in : *
19: em_mu : *
20: em_om : *
21: em_or : *
22: em_ov : *
23: em_pat : *
24: em_ph : *
25: em_pl : *
26: em_ro : *
27: em_sts : *
28: em_un : *
29: em_vl : *
30: em_hcg_hum : *
31: em_hcg_inv : *
32: em_hcg_other : *
33: em_hcg_mus : *
34: em_hcg_pin : *
35: em_hcg_rod : *
36: em_hcg_mam : *
37: em_hcg_vrt : *
38: em_sy : *
39: em_hcg_hum : *
40: em_hcg_mus : *
41: em_hcg_other : *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3761	100.0	3761	6	AX032762
2	3761	100.0	3761	8	AF228499
3	3701.4	98.4	91849	8	F6M23
4	3701.4	98.4	197975	8	ATCHRIV2
5	1755	46.7	2257	6	AX032763
6	1328	35.3	1830	8	AF228500
7	975	15.3	575	8	AY092538
8	975	15.3	575	8	AY092540
9	975	15.3	575	8	AY092541
10	975	15.3	575	8	AY092542
11	975	15.3	575	8	AY092544
12	975	15.3	575	8	AY092545
13	975	15.3	575	8	AY092546
14	975	15.3	575	8	AY092547
15	975	15.3	575	8	AY092548
16	975	15.3	575	8	AY092549
17	975	15.3	575	8	AY092552
18	975	15.3	575	8	AY092553
19	975	15.3	575	8	AY092554
20	975	15.3	575	8	AY092556
21	975	15.3	575	8	AY092557
22	975	15.3	575	8	AY092559
23	975	15.3	575	8	AY092563
24	975	15.3	575	8	AY092569
25	975	15.3	575	8	AY092570
26	975	15.3	575	8	AY092577
27	975	15.3	575	8	AY092579
28	975	15.3	575	8	AY092581
29	975	15.3	575	8	AY092583
30	975	15.3	575	8	AY092589
31	975	15.3	575	8	AY092591
32	975	15.3	575	8	AY092593
33	975	15.3	575	8	AY092594
34	975	15.3	575	8	AY092595
35	975	15.3	575	8	AY092596
36	975	15.3	575	8	AY092597
37	975	15.3	575	8	AY092598
38	975	15.3	575	8	AY092599
39	975	15.3	575	8	AY092600
40	975	15.3	575	8	AY092601
41	975	15.3	575	8	AY092602
42	975	15.3	575	8	AY092603
43	975	15.3	575	8	AY092604
44	975	15.3	575	8	AY092605
45	975	15.3	575	8	AY092606

ALIGNMENTS

RESULT 1
AX032762
LOCUS AX032762
DEFINITION Sequence 2 from Patent WO0046358.
ACCESSION AX032762
VERSION AX032762.1 GI:10279738
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 3761)
AUTHORS Dean,C., West,J. and Johanson,U.
TITLE Plant gene

JOURNAL Patent: WO 0046358-A 2 10-AUG-2000;
 DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
 JOHANSON URBAN (SE)
 FEATURES location/Qualifiers
 source 1. 3761
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 BASE COUNT 1174 a 717 c 775 g 1095 t
 ORIGIN
 Query Match 100.0%; Score 3761; DB 6; Length 3761;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3761; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTACACAGTACAGTAAACCAAGTATACCAAGATTTTATCATGGGATATCGTG 60
 DB 1 AGTACACAGTACAGTAAACCAAGTATACCAAGATTTTATCATGGGATATCGTG 60
 QY 61 TTTGAAAGCTAAAGAGACACACATCACCCCTATTAGTCAGGTAGAGACAGTAA 120
 DB 61 TTTGAAAGCTAAAGAGACACACATCACCCCTATTAGTCAGGTAGAGACAGTAA 120
 QY 121 CTTTGGGTTCATATTACGAGACAGACGTTATTGATGATGACATGTTATTAACA 180
 DB 121 CTTTGGGTTCATATTACGAGACAGACGTTATTGATGATGACATGTTATTAACA 180
 QY 181 CTGCTTAGTACTATTAAACATATATATCATGTCGTAATCATGCACTAATG 240
 DB 181 CTGCTTAGTACTATTAAACATATATATCATGTCGTAATCATGCACTAATG 240
 QY 241 TTTTCATTATCAAAATCAAGATTAAGAGAAAGTGCGTATTCATTATTGGCAT 300
 DB 241 TTTTCATTATCAAAATCAAGATTAAGAGAAAGTGCGTATTCATTATTGGCAT 300
 QY 301 AGACTCAAAAGAGTATATATATCTGACTTTTATTAATTTAACAACAATACAT 360
 DB 301 AGACTCAAAAGAGTATATATATCTGACTTTTATTAATTTAACAACAATACAT 360
 QY 361 TTTTCATTAAGCAAACTATTAAGAGCCTTAACATATATATGATTAACAGAAAGATC 420
 DB 361 TTTTCATTAAGCAAACTATTAAGAGCCTTAACATATATATGATTAACAGAAAGATC 420
 QY 421 GTTTTCCTTACCTTAAAGATAGTACTTCTTAATTAATTAATTAATTAATTAAT 480
 DB 421 GTTTTCCTTACCTTAAAGATAGTACTTCTTAATTAATTAATTAATTAATTAAT 480
 QY 481 ACATATACAGTCAATTAATTTGGTAAATTTGACGATTTAAGAGAGTGAATTAG 540
 DB 481 ACATATACAGTCAATTAATTTGGTAAATTTGACGATTTAAGAGAGTGAATTAG 540
 QY 541 GCTTTCGCAATCTTTTCTTCCGCGCAATCTCATATTCATTAATTCACCGAGTGG 600
 DB 541 GCTTTCGCAATCTTTTCTTCCGCGCAATCTCATATTCATTAATTCACCGAGTGG 600
 QY 601 GCGCAACCCACAGAGAGCGCAATCCATCTGACAGCGACATCAATCGAAGAGAG 660
 DB 601 GCGCAACCCACAGAGAGCGCAATCCATCTGACAGCGACATCAATCGAAGAGAG 660
 QY 661 AGAATATACGAAGATTTGCGAAGACAGATCAAGATTAAGATTAAGATTAAGAT 720
 DB 661 AGAATATACGAAGATTTGCGAAGACAGATCAAGATTAAGATTAAGATTAAGAT 720
 QY 721 TCTAAGAGCGCTCAATTTTGAATTCATAGAGATTAAGATTAAGATTAAGAT 780
 DB 721 TCTAAGAGCGCTCAATTTTGAATTCATAGAGATTAAGATTAAGATTAAGAT 780
 QY 781 GAAACATTCAGAGCGCAATTCGATGATCTTCAAGAGACATGATCAATGGAAGAG 840
 DB 781 GAAACATTCAGAGCGCAATTCGATGATCTTCAAGAGACATGATCAATGGAAGAG 840
 QY 841 ATTGATTCAGAGCTGAGAGTAAAGCGGTGCTCGCGCGGAGACAAATTAATTCAT 900
 DB 841 ATTGATTCAGAGCTGAGAGTAAAGCGGTGCTCGCGCGGAGACAAATTAATTCAT 900

QY 901 CAGCGATGTATTCGCTCCGCGGAAACAATGTATCTGTAGAAACACCGCTCATGTGAC 960
 DB 901 CAGCGATGTATTCGCTCCGCGGAAACAATGTATCTGTAGAAACACCGCTCATGTGAC 960
 QY 961 CAACGCTCTCAGAGATTTGACCGGAGAGCGAATTAACCGGAGGAGAGTATGCT 1020
 DB 961 CAACGCTCTCAGAGATTTGACCGGAGAGCGAATTAACCGGAGGAGAGTATGCT 1020
 QY 1021 GAGTTGATGTAGCAAGGTCTGCGTAAATATATATACCGCAATCTCTGATCAAGCT 1080
 DB 1021 GAGTTGATGTAGCAAGGTCTGCGTAAATATATATACCGCAATCTCTGATCAAGCT 1080
 QY 1081 AAGTTAATGAGAGATTTCTCAGCTTGAATTTGGCAAGAGACAGCAAGTTGTA 1140
 DB 1081 AAGTTAATGAGAGATTTCTCAGCTTGAATTTGGCAAGAGACAGCAAGTTGTA 1140
 QY 1141 TTGATGTATTTGGCAAGTTTCTTCAACAAGGCGGTAGAGCAATTAAGAGAGTGGCT 1200
 DB 1141 TTGATGTATTTGGCAAGTTTCTTCAACAAGGCGGTAGAGCAATTAAGAGAGTGGCT 1200
 QY 1201 ATGAGCTCTGCGAGACAGTTCGCTTATTAAGTGAAGTCTTTCTTCAATGCTGAT 1260
 DB 1201 ATGAGCTCTGCGAGACAGTTCGCTTATTAAGTGAAGTCTTTCTTCAATGCTGAT 1260
 QY 1261 CTTGTTAAGAGAGAGTGAAGATTTGAGAGCTTGAATTAAGATTAAGAGAGAGCTGCT 1320
 DB 1261 CTTGTTAAGAGAGAGTGAAGATTTGAGAGCTTGAATTAAGATTAAGAGAGAGCTGCT 1320
 QY 1321 GTTGCCTTGAGAGAAAGTTGATGACTGAAGAGATTAAGTGAAGTGAAGTGAATGAT 1380
 DB 1321 GTTGCCTTGAGAGAAAGTTGATGACTGAAGAGATTAAGTGAAGTGAAGTGAATGAT 1380
 QY 1381 GCAAGGCTTCTGCTTACAGTTCGCTTGTGCTTCTTCAATTAAGAGTGA 1440
 DB 1381 GCAAGGCTTCTGCTTACAGTTCGCTTGTGCTTCTTCAATTAAGAGTGA 1440
 QY 1441 GATTTGCTGATTTGATTAAGATGAGTGTGCAATGAATTAAGTGAAGTGAAGTGA 1500
 DB 1441 GATTTGCTGATTTGATTAAGATGAGTGTGCAATGAATTAAGTGAAGTGAAGTGA 1500
 QY 1501 TCAAGCTTCTTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 DB 1501 TCAAGCTTCTTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 QY 1561 ATTGCAAAAGTGTCTCTTTGTTGACATTCAGCAACATCAAGTCCATCTTGT 1620
 DB 1561 ATTGCAAAAGTGTCTCTTTGTTGACATTCAGCAACATCAAGTCCATCTTGT 1620
 QY 1621 TTTGATTAAGCTTGTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 1680
 DB 1621 TTTGATTAAGCTTGTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 1680
 QY 1681 GATTGAGCCCAATCTAGTGAATTAATCTGGAAGAGTGAAGTGAAGTGAAGTGAAG 1740
 DB 1681 GATTGAGCCCAATCTAGTGAATTAATCTGGAAGAGTGAAGTGAAGTGAAGTGAAG 1740
 QY 1741 GTTTGATTAAGCAATTTGCTGCTGTCAGAACCAATGATGATGATGATGATGATGAT 1800
 DB 1741 GTTTGATTAAGCAATTTGCTGCTGTCAGAACCAATGATGATGATGATGATGATGAT 1800
 QY 1801 GTTTCTTGTGTTTCCAAATTTGGAAGTTAAATGCTATGATGATGATGATGATGAT 1860
 DB 1801 GTTTCTTGTGTTTCCAAATTTGGAAGTTAAATGCTATGATGATGATGATGATGAT 1860
 QY 1861 TTAAGCTTAAGCAATGCTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1920
 DB 1861 TTAAGCTTAAGCAATGCTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1920
 QY 1921 GGTATAGTGAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1980
 DB 1921 GGTATAGTGAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1980

QY	1981	ACCTTTGGCATGAGGATTAAGTTTTCAGCTGCTCTAGTCTTCACTTCACTTCTTAAGATG	2040
DB	1981	ACCTTTGGCATGAGGATTAAGTTTTCAGCTGCTCTAGTCTTCACTTCACTTCTTAAGATG	2040
QY	2041	ACCAAGAGATCAATTTGAGAGGGCAAAAGGAAAGCCATGACCGCTGGCATTTGTATGA	2100
DB	2041	ACCAAGAGATCAATTTGAGAGGGCAAAAGGAAAGCCATGACCGCTGGCATTTGTATGA	2100
QY	2101	ACCCCTTCCCTTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2160
DB	2101	ACCCCTTCCCTTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2160
QY	2161	TTGATATTAATTAATTTCTCAACAGAAAGGCGCTTCAAGAGAGCTAGTGTATATAT	2220
DB	2161	TTGATATTAATTAATTTCTCAACAGAAAGGCGCTTCAAGAGAGCTAGTGTATATAT	2220
QY	2221	CAGTTATGAGTATAGAGAGCTCACAAGTTAGATCTCTGGAAGAAAGTACCAAGATGAG	2280
DB	2221	CAGTTATGAGTATAGAGAGCTCACAAGTTAGATCTCTGGAAGAAAGTACCAAGATGAG	2280
QY	2281	AGATCAAGAGCAAAATTTAGCTTGAGAGAAAGACACTCTTCACTGCAAGAGATGAG	2340
DB	2281	AGATCAAGAGCAAAATTTAGCTTGAGAGAAAGACACTCTTCACTGCAAGAGATGAG	2340
QY	2341	AAAGAAAGCAAGATCTCTAGTTAATGAGAGAAAGCGCATGCGCAAGAGATATATA	2400
DB	2341	AAAGAAAGCAAGATCTCTAGTTAATGAGAGAAAGCGCATGCGCAAGAGATATATA	2400
QY	2401	ACCAAGATTAAGAGCTGCAAGGTTGTCACCCATGGAATGCCACAGTAACTTCTTCAAT	2460
DB	2401	ACCAAGATTAAGAGCTGCAAGGTTGTCACCCATGGAATGCCACAGTAACTTCTTCAAT	2460
QY	2461	CGATTTCTCTATCTATCTAGCTGATAGAGAGCTTCTTCAAGAGAGATGACCAAGATG	2520
DB	2461	CGATTTCTCTATCTATCTAGCTGATAGAGAGCTTCTTCAAGAGAGATGACCAAGATG	2520
QY	2521	AAATATAGCTCTTGTAGATAGTAACTCTGCGCCGCTCAACATCTTCTTCCATCGCTCA	2580
DB	2521	AAATATAGCTCTTGTAGATAGTAACTCTGCGCCGCTCAACATCTTCTTCCATCGCTCA	2580
QY	2581	GAAGATCCCGGAATATATGATGTTCCACTCCACATGATGATGATGATGATGATGATG	2640
DB	2581	GAAGATCCCGGAATATATGATGTTCCACTCCACATGATGATGATGATGATGATGATG	2640
QY	2641	CATATGAACATCTGCGCCCAATTCATCTCTCCAGGTCACGAGCATAGACTTCATGAC	2700
DB	2641	CATATGAACATCTGCGCCCAATTCATCTCTCCAGGTCACGAGCATAGACTTCATGAC	2700
QY	2701	AGTACTCTCCGCTCTTGGTTTCAGGAGAGAGATGACATGACATGACATGACATGAC	2760
DB	2701	AGTACTCTCCGCTCTTGGTTTCAGGAGAGAGATGACATGACATGACATGACATGAC	2760
QY	2761	ATGAGCAACAACAGTTACCATATGATATGATATCAAAAGGTTTACAGACATTCACATCTG	2820
DB	2761	ATGAGCAACAACAGTTACCATATGATATGATATCAAAAGGTTTACAGACATTCACATCTG	2820
QY	2821	AAAGATATTTGGGTTTATCCAAATCAAAAGGTCCTCGAGTACTCATCATTAAGACCCA	2880
DB	2821	AAAGATATTTGGGTTTATCCAAATCAAAAGGTCCTCGAGTACTCATCATTAAGACCCA	2880
QY	2881	AATAGAGAGATTAATTTGTAACAAAGCTTTTGTGTTTGTGTTAAGTATGATTTAT	2940
DB	2881	AATAGAGAGATTAATTTGTAACAAAGCTTTTGTGTTTGTGTTAAGTATGATTTAT	2940
QY	2941	TTAACTCCCAACAGTCTCAAAATTAATTAATTAATTTAGTTGGGCTTAAGATCAAAATTTT	3000
DB	2941	TTAACTCCCAACAGTCTCAAAATTAATTAATTAATTTAGTTGGGCTTAAGATCAAAATTTT	3000
QY	3001	TTGCTCTCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	3060
DB	3001	TTGCTCTCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	3060
QY	3061	TCAGATACATGATATCACTATCTAATTAAGACCTCTGCTCTTGGTCTCATCTGATTTCT	3120
DB	3061	TCAGATACATGATATCACTATCTAATTAAGACCTCTGCTCTTGGTCTCATCTGATTTCT	3120
QY	3121	TCATCT	3180
DB	3121	TCATCT	3180
QY	3181	AACTTGGCTACAGTTTCCAGTACAGCAAGCTTTTCCATGATATTAATTAATTAATTA	3240
DB	3181	AACTTGGCTACAGTTTCCAGTACAGCAAGCTTTTCCATGATATTAATTAATTAATTA	3240
QY	3241	TTCTTATCTCTCTGAGAGATGATTAATTCACCGAATGCACTGAGCTTGAATTAAC	3300
DB	3241	TTCTTATCTCTCTGAGAGATGATTAATTCACCGAATGCACTGAGCTTGAATTAAC	3300
QY	3301	TCAAATCCGAACGAAACAGAGCTTTTGAATCTTTGATTTAGAGCTTTGATCTTG	3360
DB	3301	TCAAATCCGAACGAAACAGAGCTTTTGAATCTTTGATTTAGAGCTTTGATCTTG	3360
QY	3361	AGGATATCAATGAACTAGATACACTGACACTTTCAGGCTTTAACTGATTTTAAAC	3420
DB	3361	AGGATATCAATGAACTAGATACACTGACACTTTCAGGCTTTAACTGATTTTAAAC	3420
QY	3421	ATGATATGACATGATTTCCATGAGATGATGATGATGATGATGATGATGATGATGATG	3480
DB	3421	ATGATATGACATGATTTCCATGAGATGATGATGATGATGATGATGATGATGATGATG	3480
QY	3481	AATGATATGACATGATTTCCATGAGATGATGATGATGATGATGATGATGATGATGATG	3540
DB	3481	AATGATATGACATGATTTCCATGAGATGATGATGATGATGATGATGATGATGATGATG	3540
QY	3541	AAAAAAGAGACAGCTTTAGATATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3600
DB	3541	AAAAAAGAGACAGCTTTAGATATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3600
QY	3601	CATAAGTTCAGAGATGCTTAAATGCTCTATTCATCAATCAATTCATCAATTCATCA	3660
DB	3601	CATAAGTTCAGAGATGCTTAAATGCTCTATTCATCAATCAATTCATCAATTCATCA	3660
QY	3661	TAAATAGATACCAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3720
DB	3661	TAAATAGATACCAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3720
QY	3721	AGAGATTTTGTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3780
DB	3721	AGAGATTTTGTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3780
RESULT 2			
AF228499			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

Location/Qualifiers
1. .3761

mRNA

CDS

BASE COUNT
ORIGIN

Query Match	100.0%;	Score 3761;	DB 8;	Length 3761;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3761;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	AGACGACAAAGCCAACTTAAACCAAGTCACAAAGATTTATCATCGGATTAATACGG	60
Db	1	AGATCTACAAAGTCACAACTTAAACCAAGTCACAAAGATTTATCATCGGATTAATACGG	60
QY	61	TTTGAAGACTAAAAAGAGCAGCACCCATCATCCTTATAGTCAGGTAGAGTAAACAGTAA	120
Db	61	TTTGAAGACTAAAAAGAGCAGCACCCATCATCCTTATAGTCAGGTAGAGTAAACAGTAA	120
QY	121	CTTTGGGTCATATTACCGAGCAGAACCCTTATTTGTGATTAGACATGTTATTAACA	180
Db	121	CTTTGGGTCATATTACCGAGCAGAACCCTTATTTGTGATTAGACATGTTATTAACA	180
QY	181	CTGCTTTAGTGACTATTTAAACCATATATTCATGTCATATCATGCAACCTTAAGTATG	240
Db	181	CTGCTTTAGTGACTATTTAAACCATATATTCATGTCATATCATGCAACCTTAAGTATG	240
QY	241	TTTTTCATTAATCAATACAAAGATATTAAGAGAGAAAGTCGATGATTCATATTATTTGGCAT	300
Db	241	TTTTTCATTAATCAATACAAAGATATTAAGAGAGAAAGTCGATGATTCATATTATTTGGCAT	300
QY	301	AGACTCAAAAGAGTGTATATATATGACTTTTATTAATTAATTAATTAACCAATACATAT	360
Db	301	AGACTCAAAAGAGTGTATATATATGACTTTTATTAATTAATTAATTAACCAATACATAT	360
QY	361	TTTTCATTAAGCAAACTATTAAGGCCCTTAACATATATGATTTACCTCAAAAGCAAAAGTC	420
Db	361	TTTTCATTAAGCAAACTATTAAGGCCCTTAACATATATGATTTACCTCAAAAGCAAAAGTC	420
QY	421	GTTTTCTCGTACTTAAAGATAGTATCTCTAATTAATATATATTAATTTATGGAATTC	480
Db	421	GTTTTCTCGTACTTAAAGATAGTATCTCTAATTAATATATATTAATTTATGGAATTC	480
QY	481	ACATATTAACAGTTCAATATAATTTGGTAAATTTGACCCATTTAAGAGAGTGGAAATTAAG	540
Db	481	ACATATTAACAGTTCAATATAATTTGGTAAATTTGACCCATTTAAGAGAGTGGAAATTAAG	540

OY	54	GCATTCGCAATCTTTTCTTTCCTGCGGAACTCATGTCCAAATATTCACGACGGTGGCG	600
Db	541	GCATTCGCAATCTTTTCTTTCCTGCGGAACTCATGTCCAAATATTCACGACGGTGGCG	600
OY	601	GGCCAAACCCAAACGACGGGGAATCCACTGCTGACGAGCATCAATCTGAAGGAGAGA	660
Db	601	GGCCAAACCCAAACGACGGGGAATCCACTGCTGACGAGCATCAATCTGAAGGAGAGA	660
OY	661	AGAAATTAACGAGAAATTTGTGAAAACAGACTGTACAAAGTATGAGCAATTCAGATCGGTCAA	720
Db	661	AGAAATTAACGAGAAATTTGTGAAAACAGAGTGTACAAAGTATGAGCAATTCAGATCGGTCAA	720
OY	721	TCFAACACACCTCCAAATTTTTGAAATCCATAGACGAATTAAGTCCGTTTCAGTTGCAGTG	780
Db	721	TCFAACACACCTCCAAATTTTTGAAATCCATAGAGCAATTAAGTCCGTTTCAGTTGCAGTG	780
OY	781	GAAACATTAAACGGCCATTCGATGATCTTCAGAGACATTCGAGTCATTCGAAAACGA	840
Db	781	GAAACATTAAACGGCCAAATTCGATGATCTTCAGAGACATTCGAGTCATTCGAAAACGA	840
OY	841	ATTGATTCCAAACCTGAGAGAGTAACGGCGTTGTCTCTCGCCGCGGAAACAATATTTCCAT	900
Db	841	ATTGATTCCAAACCTGAGAGATTAACGGCGTGTCTCTCGCCGCGGAAACAATATTTCCAT	900
OY	901	CACCGATCTTAATTCGCTCCGCGGAAACAATGTATCTGTAAACCAACGCTCATCTGTAGC	960
Db	901	CACCGATCTTAATTCGCTCTCCGCGAACAATGTATCTGTAAACCAACGCTCATCTGTAGC	960
OY	961	CAACGCTCTCAGAGATTGTACCGGAGACGTGCAATTAACCGAGGGGGAACGTATGCT	1020
Db	961	CAACGCTCTCAGAGATTGTACCGGAGACGTGCAATTAACCGAGGGGGAACGTATGCT	1020
OY	1021	GAGTTATGTGTAGCAAAAGGCTGCGTAAATACATATACGCGAATATCTCATACACT	1080
Db	1021	GAGTTATGTGTAGCAAAAGGCTGCGTAAATACATATACGCGAATATCTCATACACT	1080
OY	1081	AAGTTAATGGAAGAGATTCCTTACGCTTTGAAATTTGGCCAAAGGAGCAGCGAAATTTGA	1140
Db	1081	AAGTTAATGGAAGAGATTCCTTACGCTTTGAAATTTGGCCAAAGGAGCAGCGAAATTTGA	1140
OY	1141	TTGGATTTGTAATGGCAAGTTTTACTTACAAAGGCGGTAGAGCAATTAATAAGAGTGGCT	1200
Db	1141	TTGGATTTGTAATGGCAAGTTTTACTTACAAAGGCGGTAGAGCAATTAATAAGAGTGGCT	1200
OY	1201	ATGAGCTCTGACAGACAAGTTTCGCTTCTATCTATCTGAGAGCTTTTCTTAAATGCTCAT	1260
Db	1201	ATGAGCTCTGACAGACAAGTTTCGCTTCTTATCTATCTGAGAGCTTTTCTTAAATGCTCAT	1260
OY	1261	CGTGTGTAAGGGAAGGTGAAGATTAAGATTTGATTTAAATATAGCGGAGACGGCTCT	1320
Db	1261	CGTGTGTAAGGGAAGGTGAAGATTAAGATTTGATTTAAATATAGCGGAGACGGCTCT	1320
OY	1321	GTTTCCTGGAGGAAAAGGTTGATGATCTGAAGGAGATTAGCTCGGCTGAGAAAATGAT	1380
Db	1321	GTTTCCTGGAGGAAAAGGTTGATGATCTGAAGGAGATTAGCTCGGCTGAGAAAATGAT	1380
OY	1381	GCAAGGGTTGCTTTTACTAGTGTCTGTTTGGTTTGGTTCCTTCAAACTTATGAGGTACA	1440
Db	1381	GCAAGGGTTGCTTTTACTAGTGTCTGTTTGGTTTGGTTCCTTCAAACTTATGAGGTACA	1440
OY	1441	GATTTGCTGATTTGTATAAGATGATGATGTTGATGATGATTTCCGGTCTTTTGAAGCG	1500
Db	1441	GATTTGCTGATTTGTATAAGATGATGATGTTGATGATGATTTCCGGTCTTTTGAAGCG	1500
OY	1501	TCACAGTTTCTTGTCCCTATGCTGACAGTGCATATATCTGTTCTACATCCGGGGAATTC	1560
Db	1501	TCACAGTTTCTTGTCCCTATGCTGACAGTGCATATATCTGTTCTACATCCGGTGAATTC	1560
OY	1561	ATTGCAAAAGTGTCTTCTTTGTGTACATCTATGCAACCAATATTAATCTTCAATCTTGT	1620
Db	1561	ATTGCAAAAGTGTCTTCTTTGTGTACATCTATGCAACCAATATTAATCTTCAATCTTGT	1620
OY	1621	TTGGAATTAAGCTTGATGTATAAATCTAGAGAGCATCAATCAATTTTATGATGCAATGACT	1680

|||||
Db 1621 TTGATTAAGCTTGATGTAATTAACATGAGACATCAATTAATTAAGAGCAATGACT 1680
OY 1681 GATTGACCAATTCAGTATGAAATTAATCTGGAAGAAAGTGAAGTCTCAACCTAG 1740
Db 1681 GATTGACCAATTCAGTATGAAATTAATCTGGAAGAAAGTGAAGTCTCAACCTAG 1740
OY 1741 GTTTGGTACGAATTTGCTTGTTCAGAACCAATTAAGTATGCTATGCTTGAATAGT 1800
Db 1741 GTTTGGTACGAATTTGCTTGTTCAGAACCAATTAAGTATGCTATGCTTGAATAGT 1800
OY 1801 GTTTCTTGCTGCTTCCAAATTTGGAAGTAAATCTGATGCTATGCTGATGCTTAA 1860
Db 1801 GTTTCTTGCTGCTTCCAAATTTGGAAGTAAATCTGATGCTATGCTGATGCTTAA 1860
OY 1861 TTAAGCTTTAAGCAATCCCACTCTAAGAAAGTGTACTTACACAAATTTCTAATGCTAT 1920
Db 1861 TTAAGCTTTAAGCAATCCCACTCTAAGAAAGTGTACTTACACAAATTTCTAATGCTAT 1920
OY 1921 GGTATGATGATCAAGTATCAAGGCTGGAATGATTAAGTAAAGCTCTGGAATGCTTAT 1980
Db 1921 GGTATGATGATCAAGTATCAAGGCTGGAATGATTAAGTAAAGCTCTGGAATGCTTAT 1980
OY 1981 ACCTTGGCATGAGAGTAAGTTCAGCTGCTAGTTCATTAATCTTAATGATG 2040
Db 1981 ACCTTGGCATGAGAGTAAGTTCAGCTGCTAGTTCATTAATCTTAATGATG 2040
OY 2041 AGCAAGAGTCAATTTGAGAGGCAAAAGCAAGGCAAGCTGCAAGCTGCAATTTGATG 2100
Db 2041 AGCAAGAGTCAATTTGAGAGGCAAAAGCAAGGCAAGCTGCAAGCTGCAATTTGATG 2100
OY 2101 ACCCTTCCCTTGACATTTATGTAACCTTTATGAACTCTTTATGATGATGCTGACCA 2160
Db 2101 ACCCTTCCCTTGACATTTATGTAACCTTTATGAACTCTTTATGATGATGCTGACCA 2160
OY 2161 TTGATATATTTATTTCTCAACAGAAAGAGCGCTTACAAAGAGCTAGCTGTATATCAT 2220
Db 2161 TTGATATATTTATTTCTCAACAGAAAGAGCGCTTACAAAGAGCTAGCTGTATATCAT 2220
OY 2221 CAGTTATGCAAGTATGAGAGTCAACAGTATGCTGCGAAAGAACTTACAGAGATGCG 2280
Db 2221 CAGTTATGCAAGTATGAGAGTCAACAGTATGCTGCGAAAGAACTTACAGAGATGCG 2280
OY 2281 AGATCAAAAGCAAAATTTGCTGAGTGAAGAAAGCACTCTTCAAGTCAAAAGATG 2340
Db 2281 AGATCAAAAGCAAAATTTGCTGAGTGAAGAAAGCACTCTTCAAGTCAAAAGATG 2340
OY 2341 AAGAGAAAGCAAGTCTCAGTTTAATGAGAAAGCGCACTTCCCAAGAAATGATA 2400
Db 2341 AAGAGAAAGCAAGTCTCAGTTTAATGAGAAAGCGCACTTCCCAAGAAATGATA 2400
OY 2401 ACCAAGCAATTAAGCTGCAAGGCTTCAACCAAGAAATGCGCACAGTAACTTCTTCAAT 2460
Db 2401 ACCAAGCAATTAAGCTGCAAGGCTTCAACCAAGAAATGCGCACAGTAACTTCTTCAAT 2460
OY 2461 CGTATTCCTATCTACCGTATGAGAGTCTTCTAGTCAAAAGAGAGCATGACCAAGATG 2520
Db 2461 CGTATTCCTATCTACCGTATGAGAGTCTTCTAGTCAAAAGAGAGCATGACCAAGATG 2520
OY 2521 AAAATATGAGCTCTGAGAGTATGCTGAGGCGGCTCAACATCTTTCTTCTATGCTCAA 2580
Db 2521 AAAATATGAGCTCTGAGAGTATGCTGAGGCGGCTCAACATCTTTCTTCTATGCTCAA 2580
OY 2581 GAAGATCCCGGAGATATATGCTTCACTTCCACATGAGTGGTTAGAGAGATGATATG 2640
Db 2581 GAAGATCCCGGAGATATATGCTTCACTTCCACATGAGTGGTTAGAGAGATGATATG 2640
OY 2641 CATATGAACATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700
Db 2641 CATATGAACATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700
OY 2701 AGTACTCTCGCTTTGGTTACGAGAGAGAGATCACTACTCTCTCCCAATTC 2760
Db 2701 AGTACTCTCGCTTTGGTTACGAGAGAGAGATCACTACTCTCTCCCAATTC 2760

Db 2701 AGTACTCTCGCTTTGGTTACGAGAGAGAGATCACTACTCTCTCCCAATTC 2760
OY 2761 ATGAGCAACAGAGTATACATATGATATGTAACAAAGGTTTACAGACATTCACATGGAAG 2820
Db 2761 ATGAGCAACAGAGTATACATATGATATGTAACAAAGGTTTACAGACATTCACATGGAAG 2820
OY 2821 AAAGATATTTGGGTTATTCATCAAAAGGCTCTCCAGTAACTCATCATTTAGACCCCA 2880
Db 2821 AAAGATATTTGGGTTATTCATCAAAAGGCTCTCCAGTAACTCATCATTTAGACCCCA 2880
OY 2881 AATAGAGGATATGTAATTTGTAACAAAGCTTTTGTGTTTGTGTTAAGTATGATATAT 2940
Db 2881 AATAGAGGATATGTAATTTGTAACAAAGCTTTTGTGTTTGTGTTAAGTATGATATAT 2940
OY 2941 TTAATCCCAACAGTCTCAAAATTTAATTAATTTGAGGCTTTAAGATGCAATTTT 3000
Db 2941 TTAATCCCAACAGTCTCAAAATTTAATTAATTTGAGGCTTTAAGATGCAATTTT 3000
OY 3001 TTGCTCTGTAATTTGATTTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 3060
Db 3001 TTGCTCTGTAATTTGATTTAAGTATGATGATGATGATGATGATGATGATGATGATGAT 3060
OY 3061 TCAATATACATCATATACATATCTAATTAAGACCTCTGCTGCTGCTGCTGCTGCTGCT 3120
Db 3061 TCAATATACATCATATACATATCTAATTAAGACCTCTGCTGCTGCTGCTGCTGCTGCT 3120
OY 3121 TCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
Db 3121 TCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
OY 3181 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
Db 3181 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
OY 3241 TTCTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
Db 3241 TTCTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
OY 3301 TCAATCCCAACAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
Db 3301 TCAATCCCAACAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
OY 3361 AGGATATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420
Db 3361 AGGATATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420
OY 3421 ATGATATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
Db 3421 ATGATATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
OY 3481 AATGATATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
Db 3481 AATGATATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
OY 3541 AAAAAGAGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
Db 3541 AAAAAGAGAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
OY 3601 CATATGATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
Db 3601 CATATGATGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
OY 3661 TAATATGATATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
Db 3661 TAATATGATATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
OY 3721 AGAGATTTTGTGAGAGGAGAGAGTTCGCGAGAGGCTTCTG 3761
Db 3721 AGAGATTTTGTGAGAGGAGAGAGTTCGCGAGAGGCTTCTG 3761

RESULT 3
FEN23

LOCUS F6N23 91849 bp DNA linear PLN 12-NOV-1999
 DEFINITION Arabidopsis thaliana BAC F6N23.
 ACCESSION AF058919
 VERSION AF058919.2 GI:6382042
 KEYWORDS Arabidopsis thaliana.
 SOURCE Arabidopsis thaliana.
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 AUTHORS Geisel, C.
 TITLE The sequence of A. thaliana F6N23
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 91849)
 AUTHORS Washington University Genome Sequencing Center.
 TITLE The A. thaliana Genome Sequencing Project.
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 91849)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-1998) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 4 (bases 1 to 91849)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (12-NOV-1999) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT Submitted by:
 Genome Sequencing Center
 Department of Genetics, Washington University,
 St. Louis, MO 63108, USA
 e-mail: twatson@watson.wustl.edu

REFERENCE
 AUTHORS Geisel, C.
 TITLE The sequence of A. thaliana F6N23
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 91849)
 AUTHORS Washington University Genome Sequencing Center.
 TITLE The A. thaliana Genome Sequencing Project.
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 91849)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-1998) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 4 (bases 1 to 91849)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (12-NOV-1999) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT Submitted by:
 Genome Sequencing Center
 Department of Genetics, Washington University,
 St. Louis, MO 63108, USA
 e-mail: twatson@watson.wustl.edu

MAPING: Clones were assigned to the YAC map by hybridization by
 M. Lodhi, Cold Spring Harbor Laboratories, and fingerprinted
 by M. Marra, Washu, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we provide a small overlap between
 neighboring submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded or sequenced with an alternate
 chemistry; an attempt was made to resolve all sequencing problems,
 such as compressions and repeats; all regions were covered by
 sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' clone is F5110, 200 bp overlap; 3' clone is F15P23, 900 bp
 overlap. Actual start of this clone is at base position 104960 of
 CEF5110; actual end is at 91149 of CEF6N23
 The clone sequenced to the left is F5110. The actual start of this
 clone is at base position 104960 of F5110; actual end is at base
 position 91149 of F6N23.

NOTES:

Coding sequences below are predicted from computer analysis, using
 the program GeneFinder (P. Green and L. Hillier, ms in preparation).

FEATURES
 Source
 Location/Qualifiers
 1..91849
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"

gene
 CDS
 /chromosome="IV"
 /map="unknown"
 /clone="F6N23"
 complement(187..907)
 /gene="F6N23.23"
 complement(join(187..347,396..512,598..907))
 /gene="F6N23.23"
 /codon_start=1
 /evidence=not_experimental
 /protein_id="AAC13624.2"
 /db_xref="GI:6382044"

gene
 CDS
 /translation="MENSNDPNDNDHOSPSRIDONDQASAVETVSTMSIDSEVPR
 TCESTSGSDIDETNSFCSEASPCNPPVITSEKSKSLGLEMOSNECVVOEISE
 PELETKERPAKLLIGEDMSGSGKVCATISNATLTLVYDNCANDSPFNVSYSW
 TSESVAVRNRESIEVEERNEISFCMLRCHIRH"
 complement(4919..7497)
 /gene="F6N23.22"
 complement(join(4919..4993,5078..5488,5567..6055,
 6157..6213,6498..6594,6935..7222,7349..7497))
 /gene="F6N23.22"
 /codon_start=1
 /evidence=not_experimental
 /protein_id="AAC13623.1"
 /db_xref="GI:3047112"

gene
 CDS
 /translation="MSITNADQVEAAAGRSKRONSLRKOLAVRVSQVATFWS
 SLTPGVLEWBGCTGDMKRRKSTESHTKGLQSKELRKILSLMBEDSGTTSY
 THNLNDDNDNCSTSMSPDLDEMYIVMSVFSQCLPGASVTGETLW
 QNDVYENKIFSRSLARVPCFPLGVLGIVTELSEDNHLNRKSLMEISAH
 OMADDEKMEIKISEKHOLPLGISDEDLHKRTISTVLNSADRSKRNDRNHRP
 NIVTSESGSFLMKQCOEOVSGFVQKKSQVYLKILHVDPLMHTFMPSPQSGIN
 ODDPSRREKENEKSVLRPMVPTVVEKESILNNTIKYLOELARVELESCGSVN
 FVRRKRTTENLNDVLTIESGNTDSTKIDDSGETEYVTRDTHLRVLEKRE
 VIVKESYRDYIVADIMETLSNLMDNFVSRSHLNKFLNLNKAFFRGAANAVSM
 IKRELRLVIGDLF"
 10235..12675
 /gene="F6N23.1"
 join(10235..10495,10693..10896,10998..11192,11341..11505,
 11633..11893,11998..12216,12315..12534,12617..12675)
 /gene="F6N23.1"
 /note="similar to the family of glycosyl hydrolases"

gene
 CDS
 /evidence=not_experimental
 /protein_id="AAC13634.1"
 /db_xref="GI:3047113"

gene
 CDS
 /translation="MAIRLNHSVIPSVYKLGAPVRSARSSILPSVGMKGVSTFSGA
 RPLVLAVKLRASTEDRVPIDDDSDQVLYVEFLHDSLGLVINDMSEVPEE
 LLDQRLTKSVNDGVVDCMVGIVESHTPOVYNSGCKLFLQMIRELGLKIQVMSF
 HEGGAVGDVHIOIPEWREIGOSNPDYFDSAGRGNTSCLMGIDKQVLRGA
 LEYFDYMRSEFERPEFEFEKIIPEIEVIGPGLGLVPSYPAQFGKRYGISEFC
 YKYLINSLKEAIEVRSCHSPGKRGPDNTEYTHPGCGPDDGDSYVGRFLW
 YSRVLIDHGDRLAANLAFESTCIAAKLSGTHMYRTASMAELTAGFYNSRDCX
 GPIDAMFKRHDALNFTCVLEKRTLDHEDFPEALADPGALWQVYNAMQNSITVASE
 NALPCYDREGYNKILENAPLTPDPGRHLSCTFYLRNPLMESQNEFERELRKH
 GEAVPDGLAPGTOETNPE"
 complement(13027..14791)
 /gene="F6N23.21"
 complement(join(13027..13335,13480..13615,13712..13840,
 13943..14001,14087..14791))
 /gene="F6N23.21"

gene
 CDS
 /evidence=not_experimental
 /protein_id="AAC13622.1"
 /db_xref="GI:3047111"
 /translation="MSILLCVPLDECYCLGCTHMLMKCLYSAGHSENMGLATSD
 FEPIPRICRLIAVEENLHDMAPPDQYGDIPNHVILKMDQOTEGRVYPIVYLD
 HENDVYLAIRGLNKLKEDYAVLNDKNGVFGVHNGLKAAMVVEEHHVYL
 BELLENNPSTLFGHSGAGVSLVLYFVONVVRGNIERRIGCFIAIPRCS
 IHLAVYADVINSVVDDELPTPTALENPKSTICPCICGCLCKDPTPEERL
 KDARCLTAPGRLLHYVRKPLRGLRRPPVVRAYVVDPRFOIYLSCHADAHYIYI
 ERSQRALDKIVKQSIIVDHEETRAIMRAASLNTIPSPSTGTFHDTBESBSA
 GSGESGSPSGMSGRMRWDQFIDCHEPVNDNSEHMLTFKQESQALLCQET"
 15539..16687
 /gene="F6N23.2"

CDS	join(15539..15554,15672..15764,15953..16026,16133..16208,16307..16353,16466..16687) /gene="F6N23.2" /note="Similar to protein kinase; coded for by A. thaliana cDNA H36947; coded for by A. thaliana cDNA H37158" /codon_start=1 /protein_id="AAC13612.1" /db_xref="GI:3047101" /translation="MTESVVEELGNVTLQLRPLSSLRKISEVVEKGYDRGDVVR ENQNDVYFLQGVRAQVYLSAGEENYQEPFLKRDPEFGHIGEDVYSADEVATE LTCLIMSDHRLLEIKSVSDSKERCLVDIILYELRDVRSVPCNPSPSPISVLMIK VTIHRYHSLSEL" 16976..19151 /gene="F6N23.3" join(16976..17111,17262..17415,17785..17889,18316..18417,18618..18666,19056..19151) /gene="F6N23.3" /codon_start=1 /evidence="not_experimental" /protein_id="AAC13635.1" /db_xref="GI:3047124" /translation="MEYSNHCPTTVASQAIAASKTVEIMKIVHNPFCFLVIGDI NPIITYDNRLKLDGNRPATRSVDAROKKRTIFLIPAFVYSCITPIETPHYESLDNRQ VATEISVPELDIRCEPNRSTKONKSPRCVAFASDLIFATISLNPFRREGMSAA LSLDSMWFHPRVRADWLLFVIVSPATSRGAFATGKMFNRKREVRCCIIEE" complement(20843..21366) /gene="F6N23.20" complement(join(20843..20944,21002..21056,21103..21155,21292..21366)) /gene="F6N23.20" /codon_start=1 /evidence="not_experimental" /protein_id="AAC13621.1" /db_xref="GI:3047110" /translation="MVAIVASVFLYNGDYGPEPNVLDPIYSSMILLVEYLIRME SYWVHLHFLMFLYIEHWIKRILAAODSVKQIFGPGEPSEKPEPDTAK" complement(22410..24753) /gene="F6N23.19" complement(join(22410..22521,23220..23538,23620..23939,24009..24137,24240..24352,24439..24548,24639..24753)) /gene="F6N23.19" /codon_start=1 /evidence="not_experimental" /protein_id="AAC13637.1" /db_xref="GI:3047126" /translation="MSPSYINVLKPPFHLMDQILTNVYKRYOGNNRKLAECLPG SPENRNDYQCOHRLKYLIDSLQKAMKKEDELSELVKYDMENDRPMKISEL RRRICKQCEPHNHLNPIILKSPYPEREILYQARGGNKRAIAKILPGRPMN IKHNNCSYKRLLEDFPSNLISGVYTGKESGPEYEPNORNTYBSCITTSIKRA KSPORDFLDJLGLMWSISSSTSLRGEESVSSVQCARLNACLETPONNNTY CVSEVEMKERLMAARFTDPISIIKTSIPASGLKRLKQYDTPETDARSHMSEE DHSVASPSKRYRVRNTGSGSKPLERRLDFLLWDEHGRNGIVNSVRLIPQKS DKSGIVRFPLR" complement(25967..27832) gene	Query Match 98.4%; Score 3701.4; DB 8; Length 91849; Best Local Similarity 99.3%; Pred. No. 0; Matches 3734; Conservative 0; Mismatches 11; Indels 16; Gaps 1;	gene
CDS	1 AGTACTACAACTCACTTAACCAAGTACACAGATTTTATCATGGGATTCGTC 60 56266 AGTACTACAACTCACTTAACCAAGTACACAGATTTTATCATGGGATTCGTC 56325 61 TTGAAGACATAAAGACACACATCCCATCTAGTGAAGTAAAGACAGTAA 120 56326 TTGAAGACATAAAGACACACATCCCATCTAGTGAAGTAAAGACAGTAA 56385 121 CTTTGGGTTCAATATACGACCAAGACCGTATTTGATTTGACATGTTAAACCA 180 56386 CTTTGGGTTCAATATACGACCAAGACCGTATTTGATTTGACATGTTAAACCA 56445 181 CTGCTTAGTGAATATTAACCAATATATACATGTCGTAATGCACTTAACATG 240 		
CDS	56446 CTGCTTAGTGAATATTAACCAATATATACATGTCGTAATGCACTTAACATG 56505 241 TTTTCATTAATCAATTAACCAAGATTAAGAGAAAAGTGCCTAGATTCATTAATGGCAT 300 56506 TTTTCATTAATCAATTAACCAAGATTAAGAGAAAAGTGCCTAGATTCATTAATGGCAT 56565 301 AGACCTAAAGAGTATATATATATCTGCTTTTATTAATTAATTAACCAATATATAT 360 56566 AGACCTAAAGAGTATATATATATCTGCTTTTATTAATTAATTAACCAATATATAT 56625 361 TTTTCATTAACCAATTAATTAAGAGTAAACATATTAATGATTAATCTCAAGAGAAAAGTGC 420 56626 TTTTCATTAACCAATTAATTAAGAGTAAACATATTAATGATTAATCTCAAGAGAAAAGTGC 56685 421 GTTTCCTCAATTAATTAAGATAGTATCTCTCAATTAATTAATTAATTAATGAACTTC 480 56686 GTTTCCTCAATTAATTAAGATAGTATCTCTCAATTAATTAATTAATTAATGAACTTC 56745 481 ACAATATACAGTTCATTAATTAATTTGATTAATTTGACAGATTTAAGAGAGTGAATTTAG 540 56746 ACAATATACAGTTCATTAATTAATTTGATTAATTTGACAGATTTAAGAGAGTGAATTTAG 56805 5686 GCTTCGCAATCTTTTCTTCCGCGCAATCTCAATGTCATTAATTAATTAATTAATTAATTAAT 56865 541 GCTTCGCAATCTTTTCTTCCGCGCAATCTCAATGTCATTAATTAATTAATTAATTAATTAAT 600 56806 GCTTCGCAATCTTTTCTTCCGCGCAATCTCAATGTCATTAATTAATTAATTAATTAATTAAT 56865 601 GCGCAATCCCAACAGCAGCGGCAATCTGTCGAGCAGATCAATCTGAACAGCAGCA 660 56866 GCGCAATCCCAACAGCAGCGGCAATCTGTCGAGCAGATCAATCTGAACAGCAGCA 56925 661 AGCAATTAACCAAGATTTGTCGAACAGATTCGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 720 56926 AGCAATTAACCAAGATTTGTCGAACAGATTCGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 56985 721 TCTAAGCACTCAATTTTGAATCAATPAGACCAATTAAGTGGCTTTCAGTTGACGTG 780 56986 TCTAAGCACTCAATTTTGAATCAATPAGACCAATTAAGTGGCTTTCAGTTGACGTG 57045 781 GAACATTAACCAAGATTTGTCGAACAGATTCGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 840 57046 GAACATTAACCAAGATTTGTCGAACAGATTCGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 57105 841 ATTGATTCCAAACTGAGAGTAAAGGCTTCTCGCGCGGGAACATTAATTTCCAT 900 57106 ATTGATTCCAAACTGAGAGTAAAGGCTTCTCGCGCGGGAACATTAATTTCCAT 57165 901 CAGCGATGTTATCGCTCCGCGGAACAATGATCTGAGAAACCAACCTGCTGAGC 960 57166 CAGCGATGTTATCGCTCCGCGGAACAATGATCTGAGAAACCAACCTGCTGAGC 57225 961 CAACCGTCTCAGAGATTTGACCGGAGACGTGAATTAACCGAGGGGAGCTATGTGT 1020 57226 CAACCGTCTCAGAGATTTGACCGGAGACGTGAATTAACCGAGGGGAGCTATGTGT 57285 1021 GAGTGTATGTTGACCAAGGCTGCGTAAATACATATACGGAATATCTGATCAAGCT 1080 57286 GAGTGTATGTTGACCAAGGCTGCGTAAATACATATACGGAATATCTGATCAAGCT 57345 1081 AAGTTAATGAGAGATTCCTTCACTTGAATTTGCCAAGAGCCAGCAAGTTGTA 1140 57346 AAGTTAATGAGAGATTCCTTCACTTGAATTTGCCAAGAGCCAGCAAGTTGTA 57405 1141 TTGGATTTGATTTGCAAGTTTACTTACAAGGGGTAAGCATTTACTTAAGAGTGGCCT 1200 57406 TTGGATTTGATTTGCAAGTTTACTTACAAGGGGTAAGCATTTACTTAAGAGTGGCCT 57465 1201 ATGAGCTTCGAGACAACTTTCCTTATTAATGAGTCTTTCTTAATGCTGTAT 1260 57466 ATGAGCTTCGAGACAACTTTCCTTATTAATGAGTCTTTCTTAATGCTGTAT 57525 1261 CTGCTTAAGGAGAGTGAAGTGAATGAGTGAATTAAGTGAAGGAGGAGAGGAGGAGTGGT 1320 57526 CTGCTTAAGGAGAGTGAAGTGAATGAGTGAATTAAGTGAAGGAGGAGAGGAGTGGT 57585 		

QY	1321	GTTCCTTGAGGAAAGAAAGCTGAGACGAGAGAGAAATTCGCGCGCTGAGAAAATGCAT	1380
Db	57586	GTTCCTTGAGGAAAGAAAGCTGAGACGAGAGAGAAATTCGCGCGCTGAGAAAATGCAT	57645
QY	1381	GCAAGGGGTTTGCTTTTACTAGTTAGTGCCTGTTTTGGTTCCTTCAACACTTATAGAGATACA	1440
Db	57646	GCAAGGGGTTTGCTTTTACTAGTTAGTGCCTGTTTTGGTTCCTTCAACACTTATAGAGATACA	57705
QY	1441	GATTTCGTGATTTGTAATGAAGATGAGGCGTTGCAATGAGATTCGCGCGCTTGAAGCGG	1500
Db	57706	GATTTCGTGATTTGTAATGAAGATGAGGCGTTGCAATGAGATTCGCGCGCTTGAAGCGG	57765
QY	1501	TCACAGTTTCTGTGCTCCATGAGTCTCAGGTACCATATTCCTGTTCTCAGTCGGTGAATTC	1560
Db	57766	TCACAGTTTCTGTGCTCCATGAGTCTCAGGTACCATATTCCTGTTCTCAGTCGGTGAATTC	57809
QY	1561	ATTGCAAAAGGATGGTTCCTTTGTTGATCATATGACCAATGACCAATTCATCTTTGTTT	1620
Db	57810	ATTGCAAAAGGATGGTTCCTTTGTTGATCATATGACCAATGACCAATTCATCTTTGTTT	57865
QY	1621	TTGCAATAGCTGTGATGATTAACACTAGAGACATCAAAATTTATAGATGCATGACT	1680
Db	57870	TTGCAATAGCTGTGATGATTAACACTAGAGACATCAAAATTTATAGATGCATGACT	57929
QY	1681	GATTGACCAAAATCCTAGCTAGAAAATTAATCTGAAAAGAACTTGAACTCAACATAG	1740
Db	57930	GATTGACCAAAATCCTAGCTAGAAAATTAATCTGAAAAGAACTTGAACTCAACATAG	57989
QY	1741	GTTTTGGTAGAAATTTGCTGTTGTCAGAACCAAAATGATAGGCTATTTGCTTCAATAGT	1800
Db	57990	GTTTTGGTAGAAATTTGCTGTTGTCAGAACCAAAATGATAGGCTATTTGCTTCAATAGT	58049
QY	1801	GTTTCTGTGGTTTCCAAATATTTGGAAATTTAAAATCGATGACTTACCTGTTGGATCTAA	1860
Db	58050	GTTTCTGTGGTTTCCAAATATTTGGAAATTTAAAATCGATGACTTACCTGTTGGATCTAA	58109
QY	1861	TTAAGCTTAAGCAATGCCAACTGCAAAAGTGGTACTTACAAATTTGATTTGGTGCATA	1920
Db	58110	TTAAGCTTAAGCAATGCCAACTGCAAAAGTGGTACTTACAAATTTGATTTGGTGCATA	58165
QY	1921	GGTATAGTTGTAATCAAGTATCAAGCGTGGATGATATTGAAAGCTCTTGAGATGTTTAT	1980
Db	58170	GGTATAGTTGTAATCAAGTATCAAGCGTGGATGATATTGAAAGCTCTTGAGATGTTTAT	58229
QY	1981	ACCTTTGGCAATGGAGGATTAAGTTTTCAGCGTCGTACCTTACTTCATCTTAAAGATG	2040
Db	58230	ACCTTTGGCAATGGAGGATTAAGTTTTCAGCGTCGTACCTTACTTCATCTTAAAGATG	58289
QY	2041	AGCAAGAGATCATTGAGAGGGCAAAACGAAAGCCAGTCACCGCTGACATTTGATGA	2100
Db	58290	AGCAAGAGATCATTGAGAGGGCAAAACGAAAGCCAGTCACCGCTGACATTTGATGA	58349
QY	2101	ACCCCTCCCTTCGACATTAATGATACCTTATGAACCTCTTATTCATCTGAGCTACACA	2160
Db	58350	ACCCCTCCCTTCGACATTAATGATACCTTATGAACCTCTTATTCATCTGAGCTACACA	58409
QY	2161	TTGATATATTTATTTTCCACACAGAAAAGACGGGTACAAAGACGTATAGCTGTTCAT	2220
Db	58410	TTGATATATTTATTTTCCACACAGAAAAGACGGGTACAAAGACGTATAGCTGTTCAT	58465
QY	2221	CAGTTATGCAAGTATGAGACCTCACAAAGTTAGATCTCTGCGAAGAACTACACAGATGAC	2280
Db	58470	CAGTTATGCAAGTATGAGACCTCACAAAGTTAGATCTCTGCGAAGAACTACACAGATGAC	58529
QY	2281	AGATCAAAAGACCAATGTTAGTCTTGAGAAAAGACATCTTCAGCTGCAAAAGAGATAG	2340
Db	58530	AGATCAAAAGACCAATGTTAGTCTTGAGAAAAGACATCTTCAGCTGCAAAAGAGATAG	58589
QY	2341	AAGAGAAAGCAAGATCTCTCAGTTTAAATGGAGAGCCGACACTTGGCCAGAGATGTATA	2400
Db	58590	AAGAGAAAGCAAGATCTCTCAGTTTAAATGGAGAGCCGACACTTGGCCAGAGATGTATA	58645

QY	2401	ACCAACAGATTAACAGCTCCAGAGTGGTCACCCATGGAATGACACAGATTCCTCAT	2460
Db	58650	ACCAACAGATTAACAGCTCCAGAGTGGTCACCCATGGAATGACACAGATTCCTCAT	5870
QY	2461	CGTATTCCTCTATCTACCGTGATAGAAAGCTTCTCTAGTCAAAAGACGATGACCAATG	2520
Db	58710	CGTATTCCTCTATCTACCGTGATAGAAAGCTTCTCTAGTCAAAAGACGATGACCAATG	5876
QY	2521	AAATATCACCTCTGTGAGTAGTATACCTGCGCCGCTACACATCTTCTCATCGCTCA	2580
Db	58770	AAATATCACCTCTGTGAGTAGTATACCTGCGCCGCTACACATCTTCTCATCGCTCA	5882
QY	2581	GAAGATCCCCGGAAATATATGTTCCACTTCCACATGTTGGGTTAGGAAGAATGTATATG	2640
Db	58830	GAAGATCCCCGGAAATATATGTTCCACTTCCACATGTTGGGTTAGGAAGAATGTATATG	5888
QY	2641	CATATGAAATCTGGCCCCCAATTCATACHTCCAGTCCAGACATAGACTTTCATGAC	2700
Db	58890	CATATGAAATCTGGCCCCCAATTCATACHTCCAGTCCAGACATAGACTTTCATGAC	5894
QY	2701	AGTACCTCCGCTTGGTTCACGAGACAGACATCCATACAGTACTCTCCCTCCAAATTC	2760
Db	58950	AGTACCTCCGCTTGGTTCACGAGACAGACATCCATACAGTACTCTCCCTCCAAATTC	5900
QY	2761	ATGACACAAACAGTTACCATATGCTATACAAAGGTTTACAGACATTCCATCTGTAAG	2820
Db	59010	ATGACACAAACAGTTACCATATGCTATACAAAGGTTTACAGACATTCCATCTGTAAG	5906
QY	2821	AAAGATATTTGGGTTATTCATCAATCAAGAGTCCCTCCGACATACATCATTTAGACCCA	2880
Db	59070	AAAGATATTTGGGTTATTCATCAATCAAGAGTCCCTCCGACATACATCATTTAGACCCA	5912
QY	2881	AAATAGGAGGAATGTAAATTTGTAAACAAAGCTTTTGTGTTTGGTTAGTATCATTTAT	2940
Db	59130	AAATAGGAGGAATGTAAATTTGTAAACAAAGCTTTTGTGTTTGGTTAGTATCATTTAT	5918
QY	2941	TTTACTCCCAACAGTCCAAAATTTAATTTAAATGTTTGGGGCTTAAGAAATGCAATTTT	3000
Db	59190	TTTACTCCCAACAGTCCAAAATTTAATTTAAATGTTTGGGGCTTAAGAAATGCAATTTT	5924
QY	3001	TTGCTCCTGTAAATGACATTTAAAGTGTAAATGTTATGCTTCAGAGGTTTTAGTCAAC	3060
Db	59250	TTGCTCCTGTAAATGACATTTAAAGTGTAAATGTTATGCTTCAGAGGTTTTAGTCAAC	5930
QY	3061	TCGAAATACCGATATACATCTATCTAAATAAAGACCTGCGCTTGGCATCTGGATTTCT	3120
Db	59310	TCGAAATACCGATATACATCTATCTAAATAAAGACCTGCGCTTGGCATCTGGATTTCT	5936
QY	3121	TCATCTCTGTCTGTCTTCCTTCTTGTCTCTGTTCACCTCTGACCAATTCGGATTC	3180
Db	59370	TCATCTCTGTCTGTCTTCCTTCTTGTCTCTGTTCACCTCTGACCAATTCGGATTC	5942
QY	3181	AACCTGTGCTTACAGTTTCCATGACACAGCTTTTCCATGAAATGTTATTTATGCGCC	3240
Db	59430	AACCTGTGCTTACAGTTTCCATGACACAGCTTTTCCATGAAATGTTATTTATGCGCC	5948
QY	3241	TTCTTTATCTTTCTTGAGGAATGAAATCAACCGAAGTCCAACTTATAGTTTACAAATCA	3300
Db	59490	TTCTTTATCTTTCTTGAGGAATGAAATCAACCGAAGTCCAACTTATAGTTTACAAATCA	5954
QY	3301	TCAAATCCGAACAGAAACAGAGCTTTTTCACATCTTTGATTTAGCAGTCTTGATCTTG	3360
Db	59550	TCAAATCCGAACAGAAACAGAGCTTTTTCACATCTTTGATTTAGCAGTCTTGATCTTG	5960
QY	3361	AGGAATATCAATGAACACTAGATACACTACACTTTCGAGGCTTTTAAACTGGATTTTAAAC	3420
Db	59610	AGGAATATCAATGAACACTAGATACACTACACTTTCGAGGCTTTTAAACTGGATTTTAAAC	5966
QY	3421	ATGAATAGAAAGATTGATTCATCGAATGAATGAGTAAAGATATGCGGAACTCTTAAACA	3480
Db	59670	ATGAATAGAAAGATTGATTCATCGAATGAATGAGTAAAGATATGCGGAACTCTTAAACA	5972
QY	3481	AATGTATGAACGGGTACGGTTCATTACAAATGTAATATACAGCACTGAGATTTATGGAAC	3540

```

Db      59730 AATGTTAGAACGGGTGGGTCTTACATGATGATTATACACACTGAGATTATGAGAG 59789
        |||
Qy      3541 AAAAAAGACACAGCTTATGATATCTACAGAGACAGACACTAAAGACAGAGAT 3600
        |||
Db      59790 AAAAAAGACACAGCTTATGATATCTACAGAGACAGACACTAAAGACAGAGAT 59849
        |||
Qy      3601 CATTAATTCAGAGATTCCTTAAATAGGCTCTATTCAATCACACATTTGGCAGACAC 3660
        |||
Db      59850 CATTAATTCAGAGATTCCTTAAATAGGCTCTATTCAATCACACATTTGGCAGACAC 59909
        |||
Qy      3661 TAATAAGATACCAAGTGGACATCGAAGAGAAATAGAGATGATATCAGAGAGAG 3720
        |||
Db      59910 TAATAAGATACCAAGTGGACATCGAAGAGAAATAGAGATGATATCAGAGAGAG 59969
        |||
Qy      3721 AGAGATTTTGGAGGAGGAGAGAGTTCGCCGAGGCTCTG 3761
        |||
Db      59970 AGAGATTTTGGAGGAGGAGAGAGTTCGCCGAGGCTCTG 60010
        |||

RESULT 4
ATCHRIV2 197975 bp DNA linear PLN 16-MAR-2000
LOCUS Arabidopsis thaliana DNA chromosome 4, contig fragment No. 2.
DEFINITION AL161472 GI:7267383
ACCESSION AL161472.2 GI:7267383
VERSION
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
            1 (bases 1 to 197975)
            Lamar, B., Stonking, T., Stumpf, J., Mewes, H.W., Lemcke, K. and
            Mayer, K.F.X.
            Unpublished
            2 (bases 1 to 197975)
            EU Arabidopsis sequencing, project.
            Direct Submission
            Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG. E-mail:
            lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@hbrsc.ac.uk
            Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATCHRIV1 at the 5' end and an
            overlap with ATCHRIV3 at the 3' end.
            Location/Qualifiers
            1. 197975
               /organism="Arabidopsis thaliana"
               /variety="Columbia"
               /db_xref="taxon:3702"
               /chromosome="4"
               /gene="AT4g00450"
               /join(10352..10384,10432..10545,10605..10709,10849..11018,
               11079..11423,11475..13674,13735..13823,13921..16530,
               16677..16862,16976..17042,17092..17547)
               /gene="AT4g00450"
               /note="Contains Aminoacyl-transfer RNA synthetases
               class-II signatures AAI343-1352; Prokaryotic membrane
               lipoprotein lipid attachment site AA349-359
               contains EST gb:AI999580.1"
               /codon_start=1
               /product="hypothetical protein"
               /protein_id="CAB80854.1"
               /db_xref="GI:7267384"
               /translation="MLSPLEVISGAVERKCLRAINESRALKKRVTKRLSDSNPROFIT
               CAEPRAGOVYVPLSGSLCKPFGPEORSCGERTKRWTESLSQHKRLRLSLADNITG
               YRRKTLFEVLIIRNVPLLRATWFIKVTYINQVHCWAINMCDIVLLFLQEHILTRHKL"

FEATURES
source
gene
CDS

```

```

CVSNGQMLNINCFMSMNPCHGTAHLLSKLTGHCRCIOPDCRRIYOHQAFTAR
KHLTLNGCINCVYSGTMLKGFPLSLILIGSSSTSPNFRFALDKEIFEIQL
LPLIYVAVLESIVLSQYVSVLVAIFARFOEPKAGSDLVDSRAVTLALIEVR
YLIVAVPEVADSEFPPLPSPVACGNVDVYSKALENKELKRSNAETLSAFOERG
YLSREFLSPTYTISTIDRSADDLAKIASAGYDPOHNKAVQALDRLASGDIIRAYS
YLFEDLCGAVDEAMITIDVSPCLSSRLMGAISTSPVCVFELIMACDGFDFAG
VPKDIKFSGRDCSOVYVLIOLKOKLIGGEFTARCKNRNNELVYSGMDVDE
SPGKHDIIVCMIOHVEYHGAGARLQLLPELIRSGIPNIVAYVOLVYSGMDVIO
PAVPERMRHRIILKQLPGCFVHEITLEBAQLRGDKLSAVRTYSNERLLRELY
EKGKIMNVLVSDOKSKITSTLSVIFPRACNAKSSKPRKHTSSVDIRLKRI
SALLOFPGMSCGVEFPVDERFONSKSVSSVSKMDQEPATGCEDECRKAPKRN
KESCYQNSNIESIASDEEDMMWIKSGKTVSSSLKVDQIEITQVPRGRKARKRN
LAOLOARIESQSGASTSHVDNVCSPHSGPVGEGNORVYVETSTSPVDVYSGN
SIKLOFDFKRSIAVWLTAVROLVEEPKSSVYVGOFNAGVVEEKTTRMLGADL
LYSLIFLIDISLIDVYAGRLVYVPRNENMCEGALVYSLRREYLLISALV
PEAMRLNRAASLMSNGKISGSALVYVRYLTKRSGSPVYVEHNNKATSEKTL
LSELDHTSGMGEYGNPLGVPAGVYNDPDIIRKITSIGARPPRVCLSRDYQARVE
EATYRLKRLIGTKAKSLAKNDGTQVAAQIVGLMDCIROTGSAOQDGLVSS
AVSALINSVGLSVARITDESIGNITYONHPSCVDSSNARYILRIHTCLLKEALGE
ROSRVFEIALATESSTALGVFAVPKSGRQGHOLSPESYDSNANNSTIDMSNGTMA
LSRATKITAASVALYIGSIHGVITLERIVGLLAKDYLDFOVFRRTKSSNGARS
MGASKVESPVEIYVHMERFLVGNCKTVSEGLVLELVGESSVVAISRMRPLPLKVEP
PASTIATVILMRPFVNSNSNSVYHEDTHRLYOSTLMAFHDVILKILPFRDPCRDQG
LYELIYABSTDAEPASVPSRGLDMHLAKVAFAIRLRLINSLIDCKVPSGSHBG
VSEAKNRHOGKTLVDKLVSVLQDLPARKHMQVBLRLLENGALAELEHNDPL
TDAIRSCPTSEKPDASENEKNTIQLITFLRLPDPVLPSEVHLFGSVDSMLK
QAEWFLAQODVLFGRKTIROKLIIVGESKGLPTPOKPMGMCNSSSDHITANKAG
KRRKEITSIIEGEVIEBGSRYVLLPRLVDENSPVGIGITTEAFVLPICIDQ
SDESRSITFENLVROFINIQQOSSLVYTRNRTSNKMGASGSEITSSKSGTSGRL
RGSPSLIRRSASNTDTPSPSPALRAASGLQIOLRLPLVIGCEPFRKTRHAL
ASTYRILGSRVYEDVAVCSPRSELKAPESYITDSSMADLSSTVLPRLRRLTAL
TRFKSLKLYCLICCKNELSRQJLDTIRNIOAMPILILSIRKSLSCOPHSPPAL
TLVPSGSTAAAGTQWRNPAISKSGTAAAGKLPKPMALPHQOQEDANDVDVPTL
LEDGTSGLSSSSNSSDMANLATCMLKGAIVRRRTDLTVGSDVDS"
/gene="AT4g00450"
/number=1
10385..10431
/gene="AT4g00450"
/number=1
10432..10545
/gene="AT4g00450"
/number=2
10546..10604
/gene="AT4g00450"
/number=2
10605..10709
/gene="AT4g00450"
/number=3
10710..10848
/gene="AT4g00450"
/number=3
10849..11018
/gene="AT4g00450"
/number=4
11019..11078
/gene="AT4g00450"
/number=4
11079..11423
/gene="AT4g00450"
/number=5
11424..11474
/gene="AT4g00450"
/number=5
11475..13674
/gene="AT4g00450"
/number=6
13675..13734
/gene="AT4g00450"
/number=6
13735..13823
/gene="AT4g00450"
/number=7

```

Intron 13824..13920
 /gene="AT4g00450"
 /number=7
 13921..16530
 /gene="AT4g00450"
 /number=8
 16531..16676
 /gene="AT4g00450"
 /number=8
 16677..16862
 /gene="AT4g00450"
 /number=9
 16863..16975
 /gene="AT4g00450"
 /number=9
 16976..17042
 /gene="AT4g00450"
 /number=10
 17043..17091
 /gene="AT4g00450"
 /number=10
 17092..17547
 /gene="AT4g00450"
 /number=11
 18027..19004
 /gene="AT4g00460"
 complement(join(18027..18146,18234..18578,18669..19004))
 /gene="AT4g00460"
 complement(join(18027..18146,18234..18578,18669..19004))
 /gene="AT4g00460"
 /note="similarity to
 contains EST gb:AI93831.1"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB80855.1"
 /db_xref="GI:7267385"
 /translation="MYSVDFPCLQLQALDSFONTEFWAEGSLSKSRSSTGSEFR
 KIVORKEKMWLPVLPSEGLSDKARKOKNREESTNOIHKAAAIINSSIISEKEI
 PSMYATLPKCGSKSVGDSIYRMGSGRPEGLDCLNISSEHVAVOIADHVEKEI
 YWRRKSLTSSKNSKMMVNDLMSTERTDKNYMAERAETLLEFCIKORPELSQSL
 DICKIOYKNDKAVLESYSHVLEGLAFNIVAWIDVLYDKTRMGSE"
 complement(18027..18146)
 /gene="AT4g00460"
 /number=1
 complement(18147..18233)
 /gene="AT4g00460"
 /number=1
 complement(18234..18578)
 /gene="AT4g00460"
 /number=2
 complement(18579..18668)
 /gene="AT4g00460"
 /number=2
 complement(18669..19004)
 /gene="AT4g00460"
 /number=3
 19070..19781
 /gene="AT4g00470"
 complement(join(19070..19221,19270..19386,19472..19781))
 /gene="AT4g00470"
 complement(join(<19070..19221,19270..19386,19472..19781))
 /note="similarity to
 contains EST gb:AA721856"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAB80856.1"
 /db_xref="GI:7267386"
 /translation="MENSINPDENDHOSPRSIDQNDQSAVEFPVSTMSIDSEVPR
 TCESTSGSDQIDETNSFCSEASPCMMVPVLTESKSSKLSGLEMSNECLVQELISE
 PLEETKEKFAKILIGEDMSGSGVCTAVTISNAITNLNDCKNDTSPDNVCSYSW
 TSEFVRAVREESIVEERNELSFICMLHCRIT"
 complement(19070..19221)
 exon

Intron /gene="AT4g00470"
 /number=1
 complement(19222..19269)
 Query Match 98.4%; Score 3701.4; DB 8; Length 197975;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 3734; Conservative 0; Mismatches 11; Indels 16; Gaps 1;
 1 AGTACACAGAGTCACACACTTAACCAAGTATACAGCAAGATTTTATCATGGATATATGCGT 60
 75140 AGTACACAGAGTCACACACTTAACCAAGTATACAGCAAGATTTTATCATGGATATATGCGT 75199
 61 TTGAGAGCTAAAAAGAGCAGACACATCACCCCATTAAGTACAGTACAGTAAAGTAA 120
 75200 TTGAGAGCTAAAAAGAGCAGACACATCACCCCATTAAGTACAGTACAGTAAAGTAA 75259
 121 CTTTGGGTCATATTACGAGCAGCAACCGTTATTGTGANTATGACATGTTATTAACCA 180
 75280 CTTTGGGTCATATTACGAGCAGCAACCGTTATTGTGANTATGACATGTTATTAACCA 75319
 181 CTGCTTATGACTATTTAAACAAATATTTACATGTCGTATATCAATGCAACCTAATG 240
 75320 CTGCTTATGACTATTTAAACAAATATTTACATGTCGTATATCAATGCAACCTAATG 75379
 241 TTTTCATTATCAATTAAGAAAGAAATTAAGAAAGTGGTATGATTCATTAATTTGGCAT 300
 75380 TTTTCATTATCAATTAAGAAAGAAATTAAGAAAGTGGTATGATTCATTAATTTGGCAT 75439
 301 AGACTCAAAAGAGTATATATATATATGACTTTTATTAATTAATTAACCAAAATACAT 360
 75440 AGACTCAAAAGAGTATATATATATATGACTTTTATTAATTAATTAACCAAAATACAT 75499
 361 TTTCATAGCAAAAGCTATAAAGCCCTTAACATATATATATGATTAATTCCTCAAGAAAGT 420
 75500 TTTCATAGCAAAAGCTATAAAGCCCTTAACATATATATATGATTAATTCCTCAAGAAAGT 75559
 421 GTTTTCTCTCTAATAAAGATAGTACTCTCTAATTAATATATATATATGTAAGTTC 480
 75560 GTTTTCTCTCTAATAAAGATAGTACTCTCTAATTAATATATATATATGTAAGTTC 75619
 481 ACATATATAGTTCATTAATAATTTGGTATTTGACCGGATTAAGAGAGTGAATTAAG 540
 75620 ACATATATAGTTCATTAATAATTTGGTATTTGACCGGATTAAGAGAGTGAATTAAG 75679
 541 GCTTCTGCAATCTTTTCTTCCGCAATCTCATATGTCATTAATTAACCGAGGTGGCG 600
 75680 GCTTCTGCAATCTTTTCTTCCGCAATCTCATATGTCATTAATTAACCGAGGTGGCG 75739
 601 GCGCAACCCACAGAGGAGGGAATCCAGTGTGAGCGAGCATCATGGAACAGAGCA 660
 75740 GCGCAACCCACAGAGGAGGGAATCCAGTGTGAGCGAGCATCATGGAACAGAGCA 75799
 661 AGAGATTTACCGAAGATTTGCAAAACAGAGTCTACAGATGATGACATTAGCATGGTCAA 720
 75800 AGAGATTTACCGAAGATTTGCAAAACAGAGTCTACAGATGATGACATTAGCATGGTCAA 75859
 721 TCTAAGCAGCTCAATTTTGAATGCAATGATGAGCAATTAAGCGGTTTTCAGTTGCAATG 780
 75860 TCTAAGCAGCTCAATTTTGAATGCAATGATGAGCAATTAAGCGGTTTTCAGTTGCAATG 75919
 781 GAAACATTTCAACGCCAATTCGATGATCTTCAAGAACACATGAGTCAATGAAACGCA 840
 75920 GAAACATTTCAACGCCAATTCGATGATCTTCAAGAACACATGAGTCAATGAAACGCA 75979
 841 ATTGATTCGAACGCGAGATACGGCGTTCGTCGCGCGCGGCAACATTAATTTTCAT 900
 75980 ATTGATTCGAACGCGAGATACGGCGTTCGTCGCGCGGCAACATTAATTTTCAT 76039
 901 CAGCGAGTGTATGCGCTCGCGGAACAAATGATCTGTAGAAACACCGTCACTGTGAGC 960
 76040 CAGCGAGTGTATGCGCTCGCGGAACAAATGATCTGTAGAAACACCGTCACTGTGAGC 76099
 961 CAACGCTCTCAGAGATTTGACCGAGAGCTGCAATTAACCGAGGAGGAGCATATGCT 1020

Db	76100	CAACCTCTCAGAGATTGTACCGGAGCGAATAAACGAGGGGGAGGTAATGCT	76159
Qy	1021	GAGTTGATGATGAGCAAGGTCGCTAAATACATATACGCAATATCTGATCAAGCT	1080
Db	76160	GATTATATGTAAGCAAGGTCGCTAAATACATATACGCAATATCTGATCAAGCT	76219
Qy	1081	AAGTAAATGAGAGATTCCTTCAAGCTTGAATTTGGCCAGAGGCGAGAGTTGTA	1140
Db	76220	AATTTAATGSAAGATTCCTTCAAGCTTGAATTTGGCCAGAGGCGAGAGTTGTA	76279
Qy	1141	TTGGAATTTATTTGGCAAGTTTACTTACAAAGGCGGTAGAGCATTTCTAAAGTCGCT	1200
Db	76280	TTGGAATTTATTTGGCAAGTTTACTTACAAAGGCGGTAGAGCATTTCTAAAGTCGCT	76339
Qy	1201	AATAGCTCTGAGAGCAAGTTTGGCTTCTATATCTGAGAGCTTTCTTCTAATGCTGAT	1260
Db	76340	ATAGCTCTGAGAGCAAGTTTGGCTTCTATATCTGAGAGCTTTCTTCTAATGCTGAT	76399
Qy	1261	CGTGTAAAGGAGAGTGAAGATTTGAGATTTAAGATGAGGCGGAGAGCGCTGCT	1320
Db	76400	CGTGTAAAGGAGAGTGAAGATTTGAGATTTAAGATGAGGCGGAGAGCGCTGCT	76459
Qy	1321	GTTCCTTGGAGAGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT	1380
Db	76460	GTTCCTTGGAGAGAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT	76519
Qy	1381	GCAGGGGGTTTGTCTTACTAGTTGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1440
Db	76520	GCAGGGGGTTTGTCTTACTAGTTGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	76579
Qy	1441	GATTTGCTGATTTGATTAAGAGAGTGTTCGATGATGATGATGATGATGATGATGATG	1500
Db	76580	GATTTGCTGATTTGATTAAGAGAGTGTTCGATGATGATGATGATGATGATGATGATG	76639
Qy	1501	TCAGAGTTTCTTGTCCCATGCTGATGATGATGATGATGATGATGATGATGATGATG	1560
Db	76640	TCAGAGTTTCTTGTCCCATGCTGATGATGATGATGATGATGATGATGATGATGATG	76683
Qy	1561	ATTGCAAAAGTGTGCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATG	1620
Db	76684	ATTGCAAAAGTGTGCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATG	76743
Qy	1621	TTTCGATTAAGCTTGTATGATTAAGTAACTAGAGAGCATCAATATTTAGAGTCAATG	1680
Db	76744	TTTCGATTAAGCTTGTATGATTAAGTAACTAGAGAGCATCAATATTTAGAGTCAATG	76803
Qy	1681	GATTGAGCCAAATCTAGTGAATTAATCTGGAAGAACTTGAAGTCAACCATAG	1740
Db	76804	GATTGAGCCAAATCTAGTGAATTAATCTGGAAGAACTTGAAGTCAACCATAG	76863
Qy	1741	GTTCGATTAAGCTTGTATGATTAAGTAACTAGAGAGCATCAATATTTAGAGTCAATG	1800
Db	76864	GTTCGATTAAGCTTGTATGATTAAGTAACTAGAGAGCATCAATATTTAGAGTCAATG	76923
Qy	1801	GTTCGATTAAGCTTGTATGATTAAGTAACTAGAGAGCATCAATATTTAGAGTCAATG	1860
Db	76924	GTTCGATTAAGCTTGTATGATTAAGTAACTAGAGAGCATCAATATTTAGAGTCAATG	76983
Qy	1861	TTAAGCTTAAGCAATGCAACTTAAGAGTGTACTTACACAAATATTTCTATGGTCA	1920
Db	76984	TTAAGCTTAAGCAATGCAACTTAAGAGTGTACTTACACAAATATTTCTATGGTCA	77043
Qy	1921	GGTATATGTAATCAAGTATCAAGGCTGATGATGATGATGATGATGATGATGATGATG	1980
Db	77044	GGTATATGTAATCAAGTATCAAGGCTGATGATGATGATGATGATGATGATGATGATG	77103
Qy	1981	ACCTTGGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2040
Db	77104	ACCTTGGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	77163
Qy	2041	ACCAAGAGTCAATTTGAGAGGCAAAACGAAAGCCAGTCAACCGCTGGCATTTGTATGA	2100
Db	77164	ACCAAGAGTCAATTTGAGAGGCAAAACGAAAGCCAGTCAACCGCTGGCATTTGTATGA	77223
Qy	2101	ACCTTCCCTTGCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2160
Db	77224	ACCTTCCCTTGCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG	77283
Qy	2161	TTGATATATTTATTTTCAACAGAAAGAGGCTCAACAGAGTCAAGTGTGTATCAT	2220
Db	77284	TTGATATATTTATTTTCAACAGAAAGAGGCTCAACAGAGTCAAGTGTGTATCAT	77343
Qy	2221	CAGTTATGAGTATGAGAGTCAAGATTTAGATGATGATGATGATGATGATGATGATGATG	2280
Db	77344	CAGTTATGAGTATGAGAGTCAAGATTTAGATGATGATGATGATGATGATGATGATGATG	77403
Qy	2281	AGATCAAGAGCAATTTGTTAGCTTGGAGAAAGACATCTTCAAGTGTGTATCAT	2340
Db	77404	AGATCAAGAGCAATTTGTTAGCTTGGAGAAAGACATCTTCAAGTGTGTATCAT	77463
Qy	2341	AAGAGAAAGCAAGATCTCTAGTTAATGAGAGAGGCGGACCTTGGCAAGAGATGATA	2400
Db	77464	AAGAGAAAGCAAGATCTCTAGTTAATGAGAGAGGCGGACCTTGGCAAGAGATGATA	77523
Qy	2401	ACCAACAGATTAAGGTCGAAAGTTGTCACCATGGAATGCCAGTAACTTCTCAT	2460
Db	77524	ACCAACAGATTAAGGTCGAAAGTTGTCACCATGGAATGCCAGTAACTTCTCAT	77583
Qy	2461	CGTATTCCTCATCTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	2520
Db	77584	CGTATTCCTCATCTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	77643
Qy	2521	AAATATCAGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2580
Db	77644	AAATATCAGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	77703
Qy	2581	GAAATCCCGGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2640
Db	77704	GAAATCCCGGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	77763
Qy	2641	CATATGACATCTGGGCGGCAATTCATCTCCAGTCAAGGAGCATATGATGATGATGATG	2700
Db	77764	CATATGACATCTGGGCGGCAATTCATCTCCAGTCAAGGAGCATATGATGATGATGATG	77823
Qy	2701	AGTACTCTCGTCTTGTGTTACGAGAGAGATGATGATGATGATGATGATGATGATGATGATG	2760
Db	77824	AGTACTCTCGTCTTGTGTTACGAGAGAGATGATGATGATGATGATGATGATGATGATGATG	77883
Qy	2761	ATGAGCAAGAGTATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2820
Db	77884	ATGAGCAAGAGTATACATATGATGATGATGATGATGATGATGATGATGATGATGATGATG	77943
Qy	2821	AAAGATATTTGGTTATCCAAATCAAGGTCCTCGAGTAACTCATCATTTAGACCCA	2880
Db	77944	AAAGATATTTGGTTATCCAAATCAAGGTCCTCGAGTAACTCATCATTTAGACCCA	78003
Qy	2881	AATAGAGAGATGATTAATTTGTAACAACTTTTGTGTTTCTTAAAGTATCATTTAT	2940
Db	78004	AATAGAGAGATGATTAATTTGTAACAACTTTTGTGTTTCTTAAAGTATCATTTAT	78063
Qy	2941	TTAAGTCCCAAGGTCGCAAAATTTAATGATGATGATGATGATGATGATGATGATGATGATG	3000
Db	78064	TTAAGTCCCAAGGTCGCAAAATTTAATGATGATGATGATGATGATGATGATGATGATGATG	78123
Qy	3001	TTGCTCCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	3060
Db	78124	TTGCTCCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	78183
Qy	3061	TCAGATATCATGATATCATATTAATATGATGATGATGATGATGATGATGATGATGATGATG	3120
Db	78184	TCAGATATCATGATATCATATTAATATGATGATGATGATGATGATGATGATGATGATGATG	78243
Qy	3121	TCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3180
Db	78244	TCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	78303

QY		318	AACCTTGNCCTTACAGTTCCCATGACACAAGCTTTCCATGAAGTATTATATGGCCGC	3242
Dd		78304	AACCTTGNCCTTACAGTTCCCATGACACAAGCTTTCCATGAAGTATTATATGGCCGC	7836
QY		3241	TTCCTTAATCTTCTTGAGAGAGATGAATTCACCGAAGATCCAACTTGAGCTTGACATCA	3300
Dd		78364	TTCCTTAATCTTCTTGAGAGAGATGAATTCACCGAAGATCCAACTTGAGAGCTTGACATCA	7842
QY		3301	TCAAAATCCGAACAGAAAACAGAGCTTTTACACTCTTTGATTATAGCAGCTTTGATCTTG	3360
Dd		78424	TCAAAATCCGAACAGAAAACAGAGCTTTTGAACATCTTTGATTATAGCAGCTTTGATCTTG	7848
QY		3361	AGGAATATCAATGAACACTAGATACACTCACACTTGACAGCTTTAAACGGAATTTAAAC	3420
Dd		78484	AGGAATATCAATGAACACTAGATACACTCACACTTGACAGCTTTAAACGGAATTTAAAC	7854
QY		3421	ATGAATTAGAAGCATTGATTCATGCGAATGTGTAACTGACATACAGCTTCTTAAACA	3480
Dd		78544	ATGAATTAGAAGCATTGATTCATGCGAATGTGTAACTGACATACAGCTTCTTAAACA	7860
QY		3481	AATGTATGAAAGCGGTAGAGGCTTCATTACAATGTGTATACAGACTGAGATTTATGGA	3540
Dd		78604	AATGTATGAAAGCGGTAGAGGCTTCATTACAATGTGTATACAGACTGAGATTTATGGA	7866
QY		3541	AAAAAAGACACAGCTTTAGATATCTPACAGAGACACAGAACCCTAAAGACAGAGAT	3600
Dd		78664	AAAAAAGACACAGCTTTAACTATCTPACAGAGACACAGAACCCTAAAGACAGAGAT	7872
QY		3601	CATAAGTTGAGAGTTCCTTAAATGAGCTCTATTTCAAACACACATTTGGCACACAACAC	3660
Dd		78724	CATAAGTTGAGAGTTCCTTAAATGAGCTCTATTTCAAATCACACATTTGGCACACAACAC	7878
QY		3661	TAAATTAAGATACCAATGGGACAAATCGAAGAGAAATTAAGATATGCATATCAGAGAGAG	3720
Dd		78784	TAAATTAAGATACCAATGGGACAAATCGAAGAGAAATTAAGATATGCATATCAGAGAGAG	7884
QY		3721	AGAGATTTTTTTGAGGAGGAGGAATTTGCCGAGAGGCTTCG	3761
Dd		78844	AGAGATTTTTTTGAGGAGGAGGAATTTGCCGAGAGGCTTCG	7884
RESULT 5				
AX032763				
LOCUS		AX032763	2257 bp	DNA linear PAT 21-SEP-2000
DEFINITION		Sequence 3 from Patent W00046358.		
ACCESSION		AX032763		
VERSION		AX032763.1	GI:10279739	
KEYWORDS				
SOURCE		.		
ORGANISM		synthetic construct.		
REFERENCE		artificial sequences.		
AUTHORS		1 (bases 1 to 2257)		
TITLE		Dean,C., West,J. and Johanson,U.		
JOURNAL		Plant gene		
		Patient: WO 0046358-A 3 10-AUG-2000;		
		DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;		
		JOHANSON URBAN (SE)		
FEATURES		Location/Qualifiers		
source		1..2257		
		/organism="synthetic construct"		
		/db_xref="Caxon:32630"		
		/note="likely CDNA sequence of the H51 FRI gene"		
BASE COUNT		698 a 443 c 493 g 623 t		
ORIGIN				
		Query Match 46.7% Score 1755; DB 6; Length 2257;		
		Best Local Similarity 82.4%; Pred. No. 0;		
		Matches 2257; Conservative 0; Mismatches 0; Indels 482; Gaps 2;		
QY		213 CATGTCGTAATCATGCAACCTPACTAGTGTTCATTATCAATATCAAGAAATAAGAGA	272	
Dd		1 CAGTCGTATATATCAACCTPACTAGTGTTCATTATCAATATCAAGAAATAAGAGA	60	

QY 1353 AGGATTAGTGGCGGTGAGAAATGATGCAAGGGGTTGCTTTACTAGTGGTCTT 1412
 |||||
 Db 1141 AGGATTAGTGGCGGTGAGAAATGATGCAAGGGGTTGCTTTACTAGTGGTCTT 1200
 |||||
 QY 1413 TGGTCTTCCTTCAAACTTTAGAGTACAGATTTGCTGATTTGATTAAGATGAGTGTTC 1472
 |||||
 Db 1201 TGGTCTTCCTTCAAACTTTAGAGTACAGATTTGCTGATTTGATTAAGATGAGTGTTC 1260
 |||||
 QY 1473 GAATGATGTCGGGCTGCTTTGAGCGGCTCACAGTTTCTTCCCTATGCTCAGGTAC 1532
 |||||
 Db 1261 GAATGATGTCGGGCTGCTTTGAGCGGCTCACAGTTTCTTCCCTATGCTCAGGTAC 1314
 |||||
 QY 1533 CAATATCTGCTCTCACTGGGTGAATTTGCAAGGTGGTCTTTGTTGATCATCAT 1592
 |||||
 Db 1315 ----- 1314
 QY 1593 CACACCAATCAATGATTCATCTTTGTTTGGATGATGATGATGATGATGATGATG 1652
 |||||
 Db 1315 ----- 1314
 QY 1653 CACATCAATATTTAGAGTACAGATGATGATGATGATGATGATGATGATGATGATG 1712
 |||||
 Db 1315 ----- 1314
 QY 1713 GGAAGAACTGGAAGCTGCAACACATGATGATGATGATGATGATGATGATGATG 1772
 |||||
 Db 1315 ----- 1314
 QY 1773 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1832
 |||||
 Db 1315 ----- 1314
 QY 1833 ATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1892
 |||||
 Db 1315 ----- 1314
 QY 1893 GTACTTACAAATATCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1952
 |||||
 Db 1315 ----- 1314
 QY 1953 GCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2012
 |||||
 Db 1348 GCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1407
 |||||
 QY 2013 TCTAGTCTTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2072
 |||||
 Db 1408 TCTAGTCTTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1467
 |||||
 QY 2073 AGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2132
 |||||
 Db 1468 AGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1489
 |||||
 QY 2133 ACTCTTATCATCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2192
 |||||
 Db 1490 ----- 1498
 QY 2193 GCTACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2252
 |||||
 Db 1499 GCTACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1558
 |||||
 QY 2253 GATCTGCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2312
 |||||
 Db 1559 GATCTGCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1618
 |||||
 QY 2313 GATCTGCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2372
 |||||
 Db 1619 GATCTGCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1678
 |||||
 QY 2373 GAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2432
 |||||
 Db 1679 GAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1738
 |||||
 QY 2433 ATGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2492

Db 1739 ATGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1798
 |||||
 QY 2493 CCTAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2552
 |||||
 Db 1799 CCTAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1858
 |||||
 QY 2553 CCGTCAACATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2612
 |||||
 Db 1859 CCGTCAACATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1918
 |||||
 QY 2613 CATGCTGCTTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2672
 |||||
 Db 1919 CATGCTGCTTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1978
 |||||
 QY 2673 CAGGTCACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2732
 |||||
 Db 1979 CAGGTCACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2038
 |||||
 QY 2733 CATCCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2792
 |||||
 Db 2039 CATCCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2098
 |||||
 QY 2793 AGGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2852
 |||||
 Db 2099 AGGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2158
 |||||
 QY 2853 CCGTCAACATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2912
 |||||
 Db 2159 CCGTCAACATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2218
 |||||
 QY 2913 TTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2951
 |||||
 Db 2219 TTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2257
 |||||
 RESULT 6
 AF228500
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana.
 Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1830)
 2 (bases 1 to 1830)
 Johnson, D., West, J., Lister, C., Michaels, S., Amasino, R. and
 Dean, C.
 Molecular analysis of FRIGIDA, a major determinant of natural
 variation in Arabidopsis flowering time
 Science 290 (5490), 344-347 (2000)
 JOURNAL MEDLINE
 PUBMED
 11030654
 REFERENCE
 2 (bases 1 to 1830)
 Johnson, D., West, J., Lister, C.
 Direct Submission
 Submitted (25-JAN-2000) Molecular Genetics, John Innes Centre,
 Colney Lane, Norwich, Norfolk NR4 7UH, U.K.
 Location/Qualifiers
 1..1830
 /organism="Arabidopsis thaliana"
 /cultivar="ecotype H51"
 /db_xref="taxon:3702"
 /chromosome="4"
 /map="between m151 and m122"
 1..1830
 /gene="FRI"
 1..1830
 /gene="FRI"
 /note="late flowering gene"

```
/codon_start=1
/product="FRIGIDA"
/protein_id="AAG23415.1"
/db_xref="gi:10801176"
/translation="MSNPPTVVAOPTTANPILIORHOSBORBRELPIVETESTSMD
ITIQSKOPFLKSIDELAFSAVETFRKROFDLQKHIESIEAIDSKESNGVYIA
IRANNFHPMLSPRRNNVSVETTVSOPSEIVFESNKEPREGRMELKSGKRLK
IVANISDAPKLMEIIPSAKLAKEPKFLVLCIEKPYLOGRAREKSPMSAROVSL
LILESFLMPDRGKVKYIESMIDEAEATAVARKRLMTEGGIAAEKMDARGLIL
VACGSPNFRSTDLDIRMSGNEITAGLRSPFLVPMVSGVSESSIRGMHIEAL
EMVTEGMEDEKPSAIVTSFLIKMSKSEFEARKRASPLAKFEATKOLAVISSVO
CMETHKIDPAKELPMQKREIVISLENDTDLQDEMEKRSLSIMEEALAKRMVNO
QIKRPLSPMEMPVTSISYSPIYKDRSPSQRDDQDELSALVSTLIGSTSPHNS
RRSPYVPLPHGGLGRSVYAEHLAPSPSPGHHRLHROYSTSLVHGQHPLOYSF
PIHGQOOLPYGIORYRHSPEERYELGISNORSRSPSNSIDPK"
```

BASE COUNT 540 a 379 c 436 g 475 t

ORIGIN

Query Match 35.3%; Score 1328; DB 8; Length 1830;
Best Local Similarity 79.2%; Pred. No. 3.3e-267;
Matches 1830; Conservative 0; Mismatches 0; Indels 482; Gaps 2;

574 ATGTCCATTATCCACCGAGGTGGCGGCAACCCACAGAGCGGCAATCCACTGCTG 633
|||||
1 ATGTCCATTATCCACCGAGGTGGCGGCAACCCACAGAGCGGCAATCCACTGCTG 60
634 CAGCAGCATCATCTGAAACAGCGAGAGAGATTACCGAAGATTGTGAAACAGAGTCT 693
|||||
61 CAGCAGCATCATCTGAAACAGCGAGAGAGATTACCGAAGATTGTGAAACAGAGTCT 120
694 ACAAGTAGGACATTCAGTCGGTCAATCTAAGCAGCTCAATTTTGAATCCATTAAC 753
|||||
121 ACAAGTAGGACATTCAGTCGGTCAATCTAAGCAGCTCAATTTTGAATCCATTAAC 180
754 GAATTCGCGCGTTTCACTGAGTGAAGAAATTCAAACGCCAATTCGATGATCTCAG 813
|||||
181 GAATTCGCGCGTTTCACTGAGTGAAGAAATTCAAACGCCAATTCGATGATCTCAG 240
814 AAGCAGCATGAGTCATCGAATAACGCAATTGATTCGAACTCGAGAGTAACGCGCTGTC 873
|||||
241 AAGCAGCATGAGTCATCGAATAACGCAATTGATTCGAACTCGAGAGTAACGCGCTGTC 300
874 CTGCGCGCGCGGAACAATAATTTCCATCAGCGGATGTTATGCCCTCCCGGGAACAATGTA 933
|||||
301 CTGCGCGCGCGGAACAATAATTTCCATCAGCGGATGTTATGCCCTCCCGGGAACAATGTA 360
934 TCTGTAGAAACCAACCGCTGAGCACAACGCTCAGAGAGATTGTACCGAGAGAGCTG 993
361 TCTGTAGAAACCAACCGCTGAGCACAACGCTCAGAGAGATTGTACCGAGAGAGCTG 420
994 AATAAACCGAGAGGAGGAGATGTGTGATGATGTAGCAAGGCTCGGTAAATAC 1053
421 AATAAACCGAGAGGAGGAGATGTGTGATGATGTAGCAAGGCTCGGTAAATAC 480
1054 AATAAACCGAGAGGAGGAGTGTGATGATGTAGCAAGGCTCGGTAAATAC 1113
|||||
481 AATAAACCGAGAGGAGGAGTGTGATGATGTAGCAAGGCTCGGTAAATAC 540
1114 TTGGCCAAAGAGCCAGAGAGTTGTATGGATTGTATGGCAAGTTTACTTACAAAGG 1173
541 TTGGCCAAAGAGCCAGAGAGTTGTATGGATTGTATGGCAAGTTTACTTACAAAGG 600
1174 CGTAGAGCATTTACTAAGAGTGCCTATGAGCTGCGAGACAATTTTGGCTTTTATA 1233
601 CGTAGAGCATTTACTAAGAGTGCCTATGAGCTGCGAGACAATTTTGGCTTTTATA 660
1234 CTGAGAGCTTTTCTTAATGCTGATCGTGTAAAGGGAAGTGAAGATTGAGATTGG 1293
661 CTGAGAGCTTTTCTTAATGCTGATCGTGTAAAGGGAAGTGAAGATTGAGATTGG 720
1294 ATTAAGATGAGGCGAGAGCGCTGCTGTTGCTTGAAGAAAGTTGATGACTGAAGA 1353
721 ATTAAGATGAGGCGAGAGCGCTGCTGTTGCTTGAAGAAAGTTGATGACTGAAGA 780

```
QY 1354 GATTTAGCTGGCGCTGAGAAAAATGATGCAAGGGGTGCTTTACTAGTTGCTGTTT 1413  
|||||  
DB 781 GATTTAGCTGGCGCTGAGAAAAATGATGCAAGGGGTGCTTTACTAGTTGCTGTTT 840  
1414 GGTGTTCTTTCAAACTTTAGAGTACAGATTGTGCTGATTTGATTAAGATGAGTGTTC 1473  
|||||  
DB 841 GGTGTTCTTTCAAACTTTAGAGTACAGATTGTGCTGATTTGATTAAGATGAGTGTTC 900  
1474 AATGAGATTGCGGCTGCTTTGAAGCGGTACAGTTTCTTGCTTCCATATGCTCAGATACC 1533  
|||||  
DB 901 AATGAGATTGCGGCTGCTTTGAAGCGGTACAGTTTCTTGCTTCCATATGCTCAGATACC 953  
1534 ATATTGCTGCTCAGCTGCTGATTTCAATTGCAAGGTTGCTTCTTTGTTGACATATC 1593  
954 ----- 953  
1594 GACCAACATCAAGTTCCATCTTTGTTTGTATGATGATTAACCTAGAGAGAGC 1653  
954 ----- 953  
1654 ACATCAAAATATTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1713  
954 ----- 953  
1714 GAAAGAACTTGAAGCTGCAACCATAGTTTGTGTAAGAAATGTTGCTTGCAGAACCA 1773  
954 ----- 953  
1774 AATGATGAGCTATTGCTTGAATAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1833  
954 ----- 953  
1834 TCGTATGACTTACCTGTTGATTAATTAAGCTTAAGCAATGCCAATCTAAGAGTGG 1893  
954 ----- 953  
1894 TACTTACAAATATTGATTTGATAGTATGATGATGATGATGATGATGATGATGATGATG 1953  
954 ----- 953  
1954 CATATTGAGCTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2013  
988 CATATTGAGCTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2047  
2014 CTAGTTTAACTTCTTAAAGATGAGCAAGAGTCAATTGAGAGGCAAAAGGAA 2073  
1048 CTAGTTTAACTTCTTAAAGATGAGCAAGAGTCAATTGAGAGGCAAAAGGAA 1107  
2074 GCCAGTACCGGCTGATTTGATGAACCTTCCCTTGACATTAATGATGATGATGATGATG 2133  
1108 GCCAGTACCGGCTGATTTGATGAACCTTCCCTTGACATTAATGATGATGATGATGATG 1128  
2134 CTCTTTATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2193  
1129 ----- 1138  
2194 CTAAAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2253  
1139 CTAAAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1198  
2254 ATCTGCGAAGAACTACAGAGATGCGATGCAAAAGCAAAATTTGATGCTTGAAGAAAG 2313  
1199 ATCTGCGAAGAACTACAGAGATGCGATGCAAAAGCAAAATTTGATGCTTGAAGAAAG 1258  
2314 ACATCTTACAGCTGAGCAAAAGATGGAAGAAAGCAAAATCTCAGATTTAATGAGAG 2273  
1259 ACATCTTACAGCTGAGCAAAAGATGGAAGAAAGCAAAATCTCAGATTTAATGAGAG 1318  
2374 AAGCGCACTTGGCAAGAGATGATTAACCAAGATTAACCAAGATTAACCAAGATTAACCA 2433  
1319 AAGCGCACTTGGCAAGAGATGATTAACCAAGATTAACCAAGATTAACCAAGATTAACCA 1378
```


OY		2434	TGGAAATGCCACCAAGTAATTCTTACTATCTGATTTCTTCATACACGTGAATGAACCTTC	2493
Db		1379	TGGAAATGCCACCAAGTAATTCTTACTATCTGATTTCTTCATACACGTGAATGAACCTTC	1438
OY		2494	CTAGTCAAAGAAGACGATGAGCAAGATGAATAATACAGTCTTGAGTAGTTACTCGGCC	2553
Db		1439	CTAGTCAAAGAAGACGATGAGCAAGATGAATAATACAGTCTTGAGTAGTTACTCGGCC	1498
OY		2554	CGTCAACATCTTTTCCATCGCTCAAGAAAGATCCCCGGAATATATGTTCCACTTCCAC	2613
Db		1499	CGTCAACATCTTTTCCATCGCTCAAGAAAGATCCCCGGAATATATGTTCCACTTCCAC	1558
OY		2614	ATGCTGGGTTAGGAAGAAGCTGTATATGACATATGACATCATGCGGCCCAATTCATCTC	2673
Db		1559	ATGCTGGGTTAGGAAGAAGCTGTATATGACATATGACATCATGCGGCCCAATTCATCTC	1618
OY		2674	CAGGTCAAGGACATAGACTTCATCGACAGTACTCTCGTCTTGTGTTACCGACAGAGAC	2733
Db		1619	CAGGTCAAGGACATAGACTTCATCGACAGTACTCTCGTCTTGTGTTACCGACAGAGAC	1678
OY		2734	ATCCACATACGATCTCTCTCCCAATTCATGAGAACAAACAAGTATACATATGGATACAA	2793
Db		1679	ATCCACATACGATCTCTCTCCCAATTCATGAGAACAAACAAGTATACATATGGATACAA	1738
OY		2794	GGGTTTACAGACATTCACACATCTGAGAAGAAAGATATTTGGGTTATCCATCAAGTCTC	2853
Db		1739	GGGTTTACAGACATTCACACATCTGAGAAGAAAGATATTTGGGTTATCCATCAAGTCTC	1798
OY		2854	CTCGCAGTACATCTCATTCATAGACCCTCAATATG	2885
Db		1799	CTCGCAGTACATCTCATTCATAGACCCTCAATATG	1830
RESULT 7				
AY092538			575 bp DNA linear PLN 03-JUN-2002	
LOCUS			Arabidopsis thaliana cultivar Lise FRIGIDA protein gene, partial cds.	
DEFINITION				
ACCESSION		AY092538		
VERSION		AY092538.1	GI:20159930	
KEYWORDS				
SOURCE			thale cress.	
ORGANISM			Arabidopsis thaliana	
			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
			Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
			Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE			1 (bases 1 to 575)	
AUTHORS			Hagenblad,J. and Nordborg,M.	
TITLE			Sequence Variation and Haplotype Structure Surrounding the Flowering Time locus Frr1 In Arabidopsis thaliana Genetics 161 (1), 289-298 (2002)	
JOURNAL				
MEDLINE			22013871	
PUBMED			12019242	
REFERENCE			2 (bases 1 to 575)	
AUTHORS			Hagenblad,J. and Nordborg,M.	
TITLE			Direct Submission	
JOURNAL			Submitted (27-MAR-2002) Department of Genetics, Lund University, Solvegatan 29, Lund SE-223 62, Sweden	
FEATURES			Location/Qualifiers	
source			1..575	
			/organism="Arabidopsis thaliana"	
			/cultivar="Lise"	
			/db_xref="taxon:3702"	
mRNA			<1..>317	
			/product-"FRIGIDA protein"	
CDS			<1..317	
			/codon_start=3	
			/product-"FRIGIDA protein"	
			/protein_id="AAI12091.1"	
			/db_xref="GI:20159931"	
			/translation="HRSRSPRYMPRLPHGLGRSVAYEHLAPNSYPGHRLHPQ YSLVTHGQRHPLQYSFPIHGQQQLPYGIQIRYRHSPDEERILGISNRRSPNSILD PK"	

ORIGIN	BASE COUNT	164 a	127 c	97 g	187 t
Query Match	15.3%	Score 575;	DB 8;	Length 575;	
Best Local Similarity	100.0%	Pred. No. 8.7e-110;			
Matches 575;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
OY	2569	CTCATCCCTCAAGAAGATCCCGGAATATATGGTTCACCTTCACATGATGGGTAGGA	2628		
DB	1	CTCATTCGCTCAAGAAGATCCCGGAATATATGGTTCACCTTCACATGATGGGTAGGA	60		
OY	2629	GAAGTATATATGATATGACATATCTGGCCCCCAATTCATCTCTCCAGATGACAGACAT	2688		
DB	61	GAAGTATATATGATATGACATATCTGGCCCCCAATTCATCTCTCCAGATGACAGACAT	120		
OY	2689	GACCTCATCGACATCTCTCCGCTTGTGGTTCACGACAGACAGACATCATCACTACT	2748		
DB	121	GACCTCATCGACATCTCTCCGCTTGTGGTTCACGACAGACAGACATCATCACTACT	180		
OY	2749	CTCTCCATTCATGACACACACATTTACATATGTTATACAAAGGTTTACAGACAT	2808		
DB	181	CTCTCCATTCATGACACACACATTTACATATGTTATACAAAGGTTTACAGACAT	240		
OY	2809	CACCATTCGAGAAAGATATTTGGGTTATCCATCAAAAGGCTCTCGAGTAACTCAT	2868		
DB	241	CACCATTCGAGAAAGATATTTGGGTTATCCATCAAAAGGCTCTCGAGTAACTCAT	300		
OY	2869	CATTAGACCCCAATAGGAGGATGTAAATTTGTAACAAAGCTTTTGTTCCTTAAG	2928		
DB	301	CATTAGACCCCAATAGGAGGATGTAAATTTGTAACAAAGCTTTTGTTCCTTAAG	360		
OY	2929	TTAGTCATTTATTTTAACTCCCAACAGCTCAAAATTTATTTATGTTTGGGCTTAGA	2988		
DB	361	TTAGTCATTTATTTTAACTCCCAACAGCTCAAAATTTATTTATGTTTGGGCTTAGA	420		
OY	2989	ATGCAAAATTTTTTCTCTCTGTATTTGACATTTAAGATGCTAAATGTTATTCCTTCAGAG	3048		
DB	421	ATGCAAAATTTTTTCTCTCTGTATTTGACATTTAAGATGCTAAATGTTATTCCTTCAGAG	480		
OY	3049	TTTTAGTCACCTCATATACATATCATATCATATTAATGACCTCTGGCTTGTGCA	3108		
DB	481	TTTTAGTCACCTCATATACATATCATATCATATTAATGACCTCTGGCTTGTGCA	540		
OY	3109	TCTGATTTCTTCATCTTCTGTCCTGTCTTC	3143		
DB	541	TCTGATTTCTTCATCTTCTGTCCTGTCTTC	575		
RESULT 8	AY092540	575 bp	DNA	linear	PLN 03-JUN-2002
LOCUS	AY092540				
DEFINITION	Arabidopsis thaliana cultivar Vilmerry FRIGIDA protein gene,				
ACCESSION	AY092540				
VERSION	AY092540.1	GI:20159934			
KEYWORDS					
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;				
TITLE	Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.				
JOURNAL	Hagenblad, J. and Nordborg, M.				
MEDLINE	Sequence Variation and Haplotype Structure Surrounding the				
POBMED	Flowering Time Locus FRI in Arabidopsis thaliana				
REFERENCE	Genetics 161 (1), 289-298 (2002)				
AUTHORS	22013871				
TITLE	12019242				
JOURNAL	2 (bases 1 to 575)				
MEDLINE	Hagenblad, J. and Nordborg, M.				
POBMED	Submitted (27-MAR-2002) Department of Genetics, Lund University,				
REFERENCE	solvegatan 29, Lund SE-223 62, Sweden				

FEATURES
source
Location/Qualifiers
1. 575
/organism="Arabidopsis thaliana"
/cultivar="Vimmerby"
/db_xref="taxon:3702"
<1. >317
/product="FRIGIDA protein"
<1. 317
/codon_start=3
/product="FRIGIDA protein"
/protein_id="AA02093.1"
/db_xref="GI:20159935"
/translation="HRSRSPETWVPLPHGIGLSYAYEHLAPNSYSPGHRLHRQ
YSPSLVHGSRHPLQYSPPIHGQOOLPYGIORYVHRSPSEERYLGLSNQSPRNSLID
PK"

BASE COUNT 164 a 127 c 97 g 187 t

ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.7e-110;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2569 CTCATCGCTCAGAGAAATCCCGGAATATATGTTCCACTTCCACATGGGGTTAGAA 2628
|||
1 CTCATCGCTCAGAGAAATCCCGGAATATATGTTCCACTTCCACATGGGGTTAGAA 60

QY 2629 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 2688
|||
61 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 120

DB 2689 GACTTCATCGACATCTACCTCCGCTTGGTTCACGACGACAGACATCCACATACAT 2748
|||
121 GACTTCATCGACATCTACCTCCGCTTGGTTCACGACGACAGACATCCACATACAT 180

QY 2749 CTCCTCCATTCATGACAAACACAGTTACCATATGTTATACAAAGGTTTACACAT 2808
|||
181 CTCCTCCATTCATGACAAACACAGTTACCATATGTTATACAAAGGTTTACACAT 240

DB 2809 CACCATCGAAGAAAGATTTGGTTTATTCATCAATCAAGGTCCTCCGAGTACAT 2868
|||
241 CACCATCGAAGAAAGATTTGGTTTATTCATCAATCAAGGTCCTCCGAGTACAT 300

QY 2869 CATATGACCCCAATATGAGAGATATTAATTTGTACAAAGCTTTTGTGTTAG 2928
|||
301 CATATGACCCCAATATGAGAGATATTAATTTGTACAAAGCTTTTGTGTTAG 360

DB 2929 TTAGTCATTTATTAACCTCCCAACATCTCAAAATTAATTAATGTTGGGCTTAA 2988
|||
361 TTAGTCATTTATTAACCTCCCAACATCTCAAAATTAATTAATGTTGGGCTTAA 420

QY 2989 ATGCAAAATTTTGTCTGCTGTAATGACATTTAAGATGCTAATGTTATGCTTAC 3048
|||
421 ATGCAAAATTTTGTCTGCTGTAATGACATTTAAGATGCTAATGTTATGCTTAC 480

DB 3049 TTTTATGCAACCTGAGATATGATATGATATGATATGATATGATATGATATGAT 3108
|||
481 TTTTATGCAACCTGAGATATGATATGATATGATATGATATGATATGATATGAT 540

QY 3109 TCTGATCTCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3143
|||
541 TCTGATCTCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575

DB

RESULT 9
LOCUS AT092541 575 bp DNA linear PLN 03-JUN-2002
DEFINITION Arabidopsis thaliana cultivar pu-2-3 FRIGIDA protein gene, partial
cde.
ACCESSION AY092541
VERSION AY092541.1 GI:20159936
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
22013871
12019242
2 (pages 1 to 575)
Hagenblad, J. and Nordborg, M.
Direct Submission
Submitted (27-MAR-2002) Department of Genetics, Lund University,
Solvegatan 29, Lund SE-223 62, Sweden
Location/Qualifiers
1. 575
/organism="Arabidopsis thaliana"
/cultivar="pu-2-3"
/db_xref="taxon:3702"
<1. >317
/product="FRIGIDA protein"
<1. 317
/codon_start=3
/product="FRIGIDA protein"
/protein_id="AA02094.1"
/db_xref="GI:20159937"
/translation="HRSRSPETWVPLPHGIGLSYAYEHLAPNSYSPGHRLHRQ
YSPSLVHGSRHPLQYSPPIHGQOOLPYGIORYVHRSPSEERYLGLSNQSPRNSLID
PK"

BASE COUNT 164 a 127 c 97 g 187 t

ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.7e-110;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2569 CTCATCGCTCAGAGAAATCCCGGAATATATGTTCCACTTCCACATGGGGTTAGAA 2628
|||
1 CTCATCGCTCAGAGAAATCCCGGAATATATGTTCCACTTCCACATGGGGTTAGAA 60

QY 2629 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 2688
|||
61 GAAGTGTATATGATATGAACATCTGGCCCAATTCATCTCCAGGTCACGACATA 120

DB 2689 GACTTCATCGACATCTACCTCCGCTTGGTTCACGACGACAGACATCCACATACAT 2748
|||
121 GACTTCATCGACATCTACCTCCGCTTGGTTCACGACGACAGACATCCACATACAT 180

QY 2749 CTCCTCCATTCATGACAAACACAGTTACCATATGTTATACAAAGGTTTACACAT 2808
|||
181 CTCCTCCATTCATGACAAACACAGTTACCATATGTTATACAAAGGTTTACACAT 240

DB 2809 CACCATCGAAGAAAGATTTGGTTTATTCATCAATCAAGGTCCTCCGAGTACAT 2868
|||
241 CACCATCGAAGAAAGATTTGGTTTATTCATCAATCAAGGTCCTCCGAGTACAT 300

QY 2869 CATATGACCCCAATATGAGAGATATTAATTTGTACAAAGCTTTTGTGTTAG 2928
|||
301 CATATGACCCCAATATGAGAGATATTAATTTGTACAAAGCTTTTGTGTTAG 360

DB 2929 TTAGTCATTTATTAACCTCCCAACATCTCAAAATTAATTAATGTTGGGCTTAA 2988
|||
361 TTAGTCATTTATTAACCTCCCAACATCTCAAAATTAATTAATGTTGGGCTTAA 420

QY 2989 ATGCAAAATTTTGTCTGCTGTAATGACATTTAAGATGCTAATGTTATGCTTAC 3048
|||
421 ATGCAAAATTTTGTCTGCTGTAATGACATTTAAGATGCTAATGTTATGCTTAC 480

DB 3049 TTTTATGCAACCTGAGATATGATATGATATGATATGATATGATATGATATGAT 3108
|||
481 TTTTATGCAACCTGAGATATGATATGATATGATATGATATGATATGATATGAT 540

QY

QY 3109 TCTGATCTCTTCATCTCTCTCTCTCTCTC 3143
|||||
Db 541 TCTGATCTCTTCATCTCTCTCTCTCTCTC 575

RESULT 10
AY092542 575 bp DNA linear PLN 03-JUN-2002
LOCUS Arabidopsis thaliana cultivar K2-9 FRIGIDA protein gene, partial
DEFINITION cds.
ACCESSION AY092542
VERSION AY092542.1 GI:20159938
KEYWORDS thale cress.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 575)
AUTHORS Hagenblad, J. and Nordborg, M.
TITLE Sequence Variation and Haplotype Structure Surrounding the Flowering Time Locus FRI in Arabidopsis thaliana
JOURNAL Genetics 161 (1), 289-298 (2002)
MEDLINE 22013871
PUBMED 12019242
REFERENCE 2 (bases 1 to 575)
AUTHORS Hagenblad, J. and Nordborg, M.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2002) Department of Genetics, Lund University, Solvegatan 29, Lund SE-223 62, Sweden
FEATURES
source location/Qualifiers
1..575
/organism="Arabidopsis thaliana"
/cultivar="K2-9"
/db_xref="taxon:3702"
<1..>317
/product="FRIGIDA protein"
<1..317
/codon_start=3
/product="FRIGIDA protein"
/protein_id="AA12095.1"
/db_xref="GI:20159938"
/translation="HRSRSPRYWVPLPHGIGRSYAYEHLAPNSYPGHRLHRO
KPSLVHGQRHPLQYSPPIHGQQLPFGIQRYRHSSEERYLGLSNRSPRNSSLD
PK"
BASE COUNT 164 a 127 c 97 g 187 t
ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.7e-110; Indels 0; Gaps 0;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2569 CTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCACATGGTGGTTAGAA 2628
|||||
Db 1 CTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCACATGGTGGTTAGAA 60

QY 2629 GAAGTGTATATGATATATACATCTGCGCCCAATTCATCTCCAGGTACGAGATA 2668
|||||
Db 61 GAAGTGTATATGATATATACATCTGCGCCCAATTCATCTCCAGGTACGAGATA 120

QY 2689 GACCTCATGACAGACTCTCGGCTTGTGTTCAAGGACAGACATCCACTACAGTACT 2748
|||||
Db 121 GACCTCATGACAGACTCTCGGCTTGTGTTCAAGGACAGACATCCACTACAGTACT 180

QY 2749 CTCCTCAATTCATGAGACACAGATTACATATGTTATACAAAGGTTTACAGACATT 2808
|||||
Db 181 CTCCTCAATTCATGAGACACAGATTACATATGTTATACAAAGGTTTACAGACATT 240

QY 2809 CACCATCTGAGAAAGATATTTGGTTTATTCATCAAGGTCTCTGCGAGTAAACAT 2868
|||||
Db 241 CACCATCTGAGAAAGATATTTGGTTTATTCATCAAGGTCTCTGCGAGTAAACAT 300

QY 2869 CATTAGACCCCAATAGAGAAATGTAATTTGTAAACAAGCTTTTGTGCTTAAG 2928
|||||

Db 301 CATAGACCCCAATAGAGAAATGTAATTTGTAAACAAGCTTTTGTGCTTAAG 360
|||||

QY 2929 TTAGTCATTTATTTAACTCCCAACAGTCTCAAAATTTATTTATGTTGGGCTTAAGA 2988
|||||

Db 361 TTAGTCATTTATTTAACTCCCAACAGTCTCAAAATTTATTTATGTTGGGCTTAAGA 420
|||||

QY 2989 ATGCAATTTTTCGCGCCGTATATGACATTTAAGATGCTAATGTTATGCTTCAGAG 3048
|||||

Db 421 ATGCAATTTTTCGCGCCGTATATGACATTTAAGATGCTAATGTTATGCTTCAGAG 480
|||||

QY 3049 TTTAGTCACCTCAGATACATCATATCATATCAATTAATAGACCTCTGCTTGGTCA 3108
|||||

Db 481 TTTAGTCACCTCAGATACATCATATCATATCAATTAATAGACCTCTGCTTGGTCA 540
|||||

QY 3109 TCTGATCTCTTCATCTCTCTCTCTCTCTC 3143
|||||
Db 541 TCTGATCTCTTCATCTCTCTCTCTCTCTC 575

RESULT 11
AY092544 575 bp DNA linear PLN 03-JUN-2002
LOCUS Arabidopsis thaliana cultivar TSO-0 FRIGIDA protein gene, partial
DEFINITION cds.
ACCESSION AY092544
VERSION AY092544.1 GI:20159942
KEYWORDS thale cress.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 575)
AUTHORS Hagenblad, J. and Nordborg, M.
TITLE Sequence Variation and Haplotype Structure Surrounding the Flowering Time Locus FRI in Arabidopsis thaliana
JOURNAL Genetics 161 (1), 289-298 (2002)
MEDLINE 22013871
PUBMED 12019242
REFERENCE 2 (bases 1 to 575)
AUTHORS Hagenblad, J. and Nordborg, M.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2002) Department of Genetics, Lund University, Solvegatan 29, Lund SE-223 62, Sweden
FEATURES
source location/Qualifiers
1..575
/organism="Arabidopsis thaliana"
/cultivar="TSO-0"
/db_xref="taxon:3702"
<1..>317
/product="FRIGIDA protein"
<1..317
/codon_start=3
/product="FRIGIDA protein"
/protein_id="AA12097.1"
/db_xref="GI:20159943"
/translation="HRSRSPRYWVPLPHGIGRSYAYEHLAPNSYPGHRLHRO
KPSLVHGQRHPLQYSPPIHGQQLPFGIQRYRHSSEERYLGLSNRSPRNSSLD
PK"
BASE COUNT 164 a 127 c 97 g 187 t
ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;
Best Local Similarity 100.0%; Pred. No. 8.7e-110; Indels 0; Gaps 0;
Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2569 CTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCACATGGTGGTTAGAA 2628
|||||

Db 1 CTCATCGCTCAAGAAGATCCCGGAATATATGTTCCACTTCCACATGGTGGTTAGAA 60
|||||

QY 2629 GAAGTGTATATGATATATACATCTGCGCCCAATTCATCTCCAGGTACGAGATA 2668
|||||

Db 61 GAAGTGTATATGATATGACATCTGGCCCAATTCATCTCCAGGTACAGGACATA 120
 QY 2689 GACTTCATCGACAGTACTCTCCGCTTTGGTTTCAGGACAGACATCCACTACAGTACT 2748
 Db 121 GACTTCATCGACAGTACTCTCCGCTTTGGTTTCAGGACAGACATCCACTACAGTACT 180
 QY 2749 CTCCTCCATTCATGAGACAAACAGTTTCCATATGCTTACAAAGGTTTACAGACATT 2808
 Db 181 CTCCTCCATTCATGAGACAAACAGTTTCCATATGCTTACAAAGGTTTACAGACATT 240
 QY 2809 CACCATCTGAAGAAAGATATTTGGGTTTATCCATCAAGAGTCTCTCGACATACAT 2868
 Db 241 CACCATCTGAAGAAAGATATTTGGGTTTATCCATCAAGAGTCTCTCGACATACAT 300
 QY 2869 CATTAAGCCCCAATATGAGAGAAATGTAATTTGTAAACAAAGCTTTTGTGGCTTAAG 2928
 Db 301 CATTAAGCCCCAATATGAGAGAAATGTAATTTGTAAACAAAGCTTTTGTGGCTTAAG 360
 QY 2929 TTATGATTTATTTAACTCCCAACAGTCTCAAAATTTAATATGTTTGGGCTTAAGA 2988
 Db 361 TTATGATTTATTTAACTCCCAACAGTCTCAAAATTTAATATGTTTGGGCTTAAGA 420
 QY 2989 ATGCAAAATTTTGTCTCTGCTTAATGACATTTAAGATGCTAATGTTATGCTTACAGG 3048
 Db 421 ATGCAAAATTTTGTCTCTGCTTAATGACATTTAAGATGCTAATGTTATGCTTACAGG 480
 QY 3049 TTTTATGCAACCTGACATATGATATGATCTATCTAATATAGACCTTGCTGGCTTAAGA 3108
 Db 481 TTTTATGCAACCTGACATATGATATGATCTATCTAATATAGACCTTGCTGGCTTAAGA 540
 QY 3109 TCTGATTTCTCTCATCTCTGCTGCTGCTGCTTC 3143
 Db 541 TCTGATTTCTCTCATCTCTGCTGCTGCTTC 575

RESULT 12
 AY092545 575 bp DNA linear PLN 03-JUN-2002
 LOCUS Arabidopsis thaliana cultivar Ler FRIGIDA protein gene, partial
 DEFINITION cds.
 ACCESSION AY092545 GI:20159944
 VERSION AY092545.1
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana
 thale cress.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 1 (bases 1 to 575)
 Hagenblad, J. and Nordborg, M.
 Sequence Variation and Haplotype Structure Surrounding the
 Flowering Time locus FRI in Arabidopsis thaliana
 Genetics 161 (1), 289-298 (2002)
 JOURNAL
 MEDLINE 12019242
 PUBMED 22013871
 2 (bases 1 to 575)
 Hagenblad, J. and Nordborg, M.
 Direct Submission
 TITLE Submitted (27-MAR-2002) Department of Genetics, Lund University,
 Solvegatan 29, Lund SE-223 62, Sweden
 JOURNAL
 FEATURES
 source
 1..575
 /organism="Arabidopsis thaliana"
 /cultivar="Ler"
 /db_xref="taxon:3702"
 <1..>317
 /product="FRIGIDA protein"
 <1..317
 /codon_start=3
 /product="FRIGIDA protein"
 /protein_id="AAM12098.1"
 /db_xref="gi:20159945"
 /translation="HRSRSPBYMPLPHGLGSYAYEHLAPNSYSPGHRLHRO

YPSLIVHGQRHPLQISPPILHGQOULPYGIQIRYRHSPEERTLGLSNRSPRNSSLD
 PK*
 BASE COUNT 164 a 127 c 97 g 187 t
 ORIGIN
 Query Match 15.3%; Score 575; DB 8; Length 575;
 Best Local Similarly 100.0%; Pred. No. 8.7e-110;
 Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2569 CTCATGCTCAGAGAGATCCCGGAATATGCTTCCACTTCCATAGTGGGTAGGAA 2628
 Db 1 CTCATGCTCAGAGAGATCCCGGAATATGCTTCCACTTCCATAGTGGGTAGGAA 60
 QY 2629 GAAGTGTATATGATATGAACATCTGGCCCCCAATTCATCTCTCCAGTCCAGGACATA 2688
 Db 61 GAAGTGTATATGATATGAACATCTGGCCCCCAATTCATCTCTCCAGTCCAGGACATA 120
 QY 2689 GACTTCATCGACAGTACTCTCCGCTTTGGTTTACGACAGACATCCACTACAGTACT 2748
 Db 121 GACTTCATCGACAGTACTCTCCGCTTTGGTTTACGACAGACATCCACTACAGTACT 180
 QY 2749 CTCCTCCAATTCATGACACAACAGTTTACCATATGATATGATTAAGAGGTTTACAGACATT 2808
 Db 181 CTCCTCCAATTCATGACACAACAGTTTACCATATGATATGATTAAGAGGTTTACAGACATT 240
 QY 2809 CACCATCTGAAGAAAGATATTTGGGTTTATCCATCAAGAGTCTCTCGACATACAT 2868
 Db 241 CACCATCTGAAGAAAGATATTTGGGTTTATCCATCAAGAGTCTCTCGACATACAT 300
 QY 2869 CATTAAGCCCCAATATGAGAGAAATGTAATTTGTAAACAAAGCTTTTGTGGCTTAAG 2928
 Db 301 CATTAAGCCCCAATATGAGAGAAATGTAATTTGTAAACAAAGCTTTTGTGGCTTAAG 360
 QY 2929 TTATGATTTATTTAACTCCCAACAGTCTCAAAATTTAATATGTTTGGGCTTAAGA 2988
 Db 361 TTATGATTTATTTAACTCCCAACAGTCTCAAAATTTAATATGTTTGGGCTTAAGA 420
 QY 2989 ATGCAAAATTTTGTCTCTGCTTAATGACATTTAAGATGCTAATGTTATGCTTACAGG 3048
 Db 421 ATGCAAAATTTTGTCTCTGCTTAATGACATTTAAGATGCTAATGTTATGCTTACAGG 480
 QY 3049 TTTTATGCAACCTGACATATGATATGATCTATCTAATATAGACCTTGCTGGCTTAAGA 3108
 Db 481 TTTTATGCAACCTGACATATGATATGATCTATCTAATATAGACCTTGCTGGCTTAAGA 540
 QY 3109 TCTGATTTCTCTCATCTCTGCTGCTGCTGCTTC 3143
 Db 541 TCTGATTTCTCTCATCTCTGCTGCTGCTTC 575
 RESULT 13
 AY092546 575 bp DNA linear PLN 03-JUN-2002
 LOCUS Arabidopsis thaliana cultivar Kondara FRIGIDA protein gene, partial
 DEFINITION cds.
 ACCESSION AY092546 GI:20159946
 VERSION AY092546.1
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana
 thale cress.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 1 (bases 1 to 575)
 Hagenblad, J. and Nordborg, M.
 Sequence Variation and Haplotype Structure Surrounding the
 Flowering Time locus FRI in Arabidopsis thaliana
 Genetics 161 (1), 289-298 (2002)
 JOURNAL
 MEDLINE 22013871
 PUBMED 12019242
 2 (bases 1 to 575)
 Hagenblad, J. and Nordborg, M.
 Direct Submission
 TITLE

[REDACTED]

Db 481 TTTAGTCAACCTGATACATGATATACACTATCAATATAGACCTTGCTGTGCTCA 540

QY 3109 TCTGATTCCTTCATCTCTGCTGCTGCTTC 3143

Db 541 TCTGATTCCTTCATCTCTGCTGCTGCTTC 575

RESULT 15

AY092548 575 bp DNA linear PLN 03-JUN-2002

LOCUS Arabidopsis thaliana cultivar TAMM-46 FRIGIDA protein gene, partial cds.

ACCESSION AY092548

VERSION AY092548.1 GI:20159950

KEYWORDS

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 575)

AUTHORS Hagenblad, J. and Nordborg, M.

TITLE Sequence Variation and Haplotype Structure Surrounding the Flowering Time Locus FRI in Arabidopsis thaliana

JOURNAL Genetics 161 (1), 289-298 (2002)

MEDLINE 12019242

PMID 12013871

REFERENCE 2 (bases 1 to 575)

AUTHORS Hagenblad, J. and Nordborg, M.

TITLE Direct Submission

JOURNAL Submitted (27-MAR-2002) Department of Genetics, Lund University, Solvegatan 29, Lund SE-223 62, Sweden

FEATURES

source

1. 575

/organism="Arabidopsis thaliana"

/cultivar="TAMM-46"

/db_xref="taxon:3702"

<1..>317

/product="FRIGIDA protein"

<1..317

/codon_start=3

/product="FRIGIDA protein"

/protein_id="AA12101.1"

/db_xref="GI:20159951"

/translation="HRSRSPERYWPLPHGLGRSYAYEHLAPNSYSPGHRLHRQ

ISPSLVHGRHPLDYSPIHQQDLPYGIQRYRHSPSEERYLGLSNRSPPNSSLD

PK"

BASE COUNT 164 a 127 c 97 g 187 t

ORIGIN

Query Match 15.3%; Score 575; DB 8; Length 575;

Best Local Similarity 100.0%; Pred. No. 8.7e-110; Indels 0; Gaps 0;

Matches 575; Conservative 0; Mismatches 0;

QY 2569 CTCATCGCTCAAGAGATCCCGGAAATATATGTTCCATCCATGATGGTTAGAA 2628

Db 1 CTCATCGCTCAAGAGATCCCGGAAATATATGTTCCATCCATGATGGTTAGAA 60

QY 2629 GAAGGTATATGCAATATGCAATGTCGCCCCAAATTCATCTCCAGGTACAGACATA 2688

Db 61 GAAGGTATATGCAATATGCAATGTCGCCCCAAATTCATCTCCAGGTACAGACATA 120

QY 2689 GACTTCATGACAGTACTCTCCGCTTGGTTGTTACGAGACAGACATCCACTACAGTACT 2748

Db 121 GACTTCATGACAGTACTCTCCGCTTGGTTGTTACGAGACAGACATCCACTACAGTACT 180

QY 2749 CTCCTCCATTCATGAGCAACACAGTTACATATGATATCAAAAGGTTTACAGACATT 2808

Db 181 CTCCTCCATTCATGAGCAACACAGTTACATATGATATCAAAAGGTTTACAGACATT 240

QY 2809 CACCATCTGAAGAAAGATATTTGGTTATTCATCAAGGTCCTCCGACGATCACTCAT 2868

Db 241 CACCATCTGAAGAAAGATATTTGGTTATTCATCAAGGTCCTCCGACGATCACTCAT 300

QY 2869 CATTAGACCCCAATAGAGAGATGTAATTTGTAAACAAGCTTTTGTGCTTAAG 2928

Db 301 CATTAGACCCCAATAGAGAGATGTAATTTGTAAACAAGCTTTTGTGCTTAAG 360

QY 2929 TTACTCATTTTAACTCCCAACAGCTCCAAATTTAATTTAATGTTGGGGCTTAAGA 2988

Db 361 TTACTCATTTTAACTCCCAACAGCTCCCAATTTAATTTAATGTTGGGGCTTAAGA 420

QY 2989 ATGCAATTTTGTGCTCTGTATTTGACATTTAAGATGTAATTTATGCTTCAGAGG 3048

Db 421 ATGCAATTTTGTGCTCTGTATTTGACATTTAAGATGTAATTTATGCTTCAGAGG 480

QY 3049 TTTTAGTCAACCTCAGATACATGATATCATATGTAATATAGACCTCGGCTGTGTC 3108

Db 481 TTTTAGTCAACCTCAGATACATGATATCATATGTAATATAGACCTCGGCTGTGTC 540

QY 3109 TCTGATTCCTTCATCTCTGCTGCTGCTGCTTC 3143

Db 541 TCTGATTCCTTCATCTCTGCTGCTGCTTC 575

Search completed: July 11, 2003, 13:21:49

Job time : 9430 secs

XX Claim 4: Fig 4: 73pp: English.

CC The present sequence represents the (late flowering) H51 genomic
 CC sequence of the first 17 kb of cosmid B4m13 encompassing FRI (one
 CC locus-PRGIDA) locus of Arabidopsis. The FRI gene encodes a
 CC polypeptide capable of specifically altering the flowering time of a
 CC plant. The FRI polynucleotide is used to transform plants, so that
 CC the flowering time of a plant is altered. This is used, for example,
 CC for plants in which the leaves or tubers are a commercial product,
 CC where it is desirable to avoid 'blotting' (initiation of flowers and
 CC stem elongation) at too early a stage. Conversely, it may be desirable
 CC to alter flowering under certain circumstances e.g. to vary flower
 CC production across the seasons.

XX Sequence 3761 BP: 1174 A; 717 C; 775 G; 1095 T; 0 other;

Query Match 100.0%; Score 3761; DB 21; Length 3761;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3761; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTACTCAAGTCACTTAACCAAGTACACAGGATTATTCATGGGATTATCGTG 60
 DB 1 AGTACTCAAGTCACTTAACCAAGTACACAGGATTATTCATGGGATTATCGTG 60
 QY 61 TTGGAAGCTAAAGAGACACACATCCATCCCATTTAGTCAGTAGAGTAAGCA 120
 DB 61 TTGGAAGCTAAAGAGACACACATCCATCCCATTTAGTCAGTAGAGTAAGCA 120
 QY 121 CTTTGGGCTCATTTATTCAGAGCAAGACGTTATTTGTGATTTAGCATGTTAAACCA 180
 DB 121 CTTTGGGCTCATTTATTCAGAGCAAGACGTTATTTGTGATTTAGCATGTTAAACCA 180
 QY 181 CTGCTTTAGTGAATTTTAAACCAATATTTACATGTCGTATCATGCAACCTACTAG 240
 DB 181 CTGCTTTAGTGAATTTTAAACCAATATTTACATGTCGTATCATGCAACCTACTAG 240
 QY 241 TTTTCAATTAATCAATTAAGAAATTAAGAGAAAGCGGTAGTTCAATTAATTTGGCAT 300
 DB 241 TTTTCAATTAATCAATTAAGAAATTAAGAGAAAGCGGTAGTTCAATTAATTTGGCAT 300
 QY 301 AGACTCAAAAGAGTATATATATCTGACTTTTAAATTTTAAACACAAATACATAT 360
 DB 301 AGACTCAAAAGAGTATATATATCTGACTTTTAAATTTTAAACACAAATACATAT 360
 QY 361 TTTTCAATTAAGCAAAATATTAAGAGCGCTTAACATATATGATTACCTAAAGGAAAGATC 420
 DB 361 TTTTCAATTAAGCAAAATATTAAGAGCGCTTAACATATATGATTACCTAAAGGAAAGATC 420
 QY 421 GTTTCTCTACTTAAAGATAGTTACTTCTTAATTAATTAATTAATTTATGTAACCTTC 480
 DB 421 GTTTCTCTACTTAAAGATAGTTACTTCTTAATTAATTAATTAATTTATGTAACCTTC 480
 QY 481 ACAATATACAGTTCAATTAATTTGTAATTTGACCGAATTTAAGAGAGGGAATTAGG 540
 DB 481 ACAATATACAGTTCAATTAATTTGTAATTTGACCGAATTTAAGAGAGGGAATTAGG 540
 QY 541 GCTTGTCAATCTTTTCTTCGCGCAATCTCAATGTCATATTCACACGAGCGTGGCG 600
 DB 541 GCTTGTCAATCTTTTCTTCGCGCAATCTCAATGTCATATTCACACGAGCGTGGCG 600
 QY 601 GGGCAACCCAGCAAGAGCGGCAATCTCAATGTCAGAGGACATCAATTTGAACACGAGCA 660
 DB 601 GGGCAACCCAGCAAGAGCGGCAATCTCAATGTCAGAGGACATCAATTTGAACACGAGCA 660
 QY 661 AGAGAATTAACGAAGATTGTGAAACAGAGTCTCAACAGTATGACATTAGCATGCTGCA 720
 DB 661 AGAGAATTAACGAAGATTGTGAAACAGAGTCTCAACAGTATGACATTAGCATGCTGCA 720
 QY 721 TCTAAGCAGCTCAATTTTGAATTCATATGAGCAATTTAGCTGCTTTCAGTTGCGAGTG 780
 DB 721 TCTAAGCAGCTCAATTTTGAATTCATATGAGCAATTTAGCTGCTTTCAGTTGCGAGTG 780

QY 781 GAACATTCAAAGCCCAATTCGATGATCTTCAGAAAGCACATCGAGTCAATCGAAAAACCA 840
 DB 781 GAACATTCAAAGCCCAATTCGATGATCTTCAGAAAGCACATCGAGTCAATCGAAAAACCA 840
 QY 841 ATTGATTCCAAACTCGAGAGTAAGCGGCTTTCCTCGCCGCGGCAACATATTTCCAT 900
 DB 841 ATTGATTCCAAACTCGAGAGTAAGCGGCTTTCCTCGCCGCGGCAACATATTTCCAT 900
 QY 901 CAGCCGATGTATTCGCTCGCGGGAACAATGATCTGTGAAGAACACCGCTCATGTGAGC 960
 DB 901 CAGCCGATGTATTCGCTCGCGGGAACAATGATCTGTGAAGAACACCGCTCATGTGAGC 960
 QY 961 CAACCGTCTCAGAGATTTGACCGAGACGTGCAATTAACCGAGGCGGAGCATATGTGT 1020
 DB 961 CAACCGTCTCAGAGATTTGACCGAGACGTGCAATTAACCGAGGCGGAGCATATGTGT 1020
 QY 1021 GAGTTGATGTGTGCAAGGCTCGCTTAATTCATTTATGCGAATATCTGATCACT 1080
 DB 1021 GAGTTGATGTGTGCAAGGCTCGCTTAATTCATTTATGCGAATATCTGATCACT 1080
 QY 1081 AAGTTAATGAGAGATTTCCCTCAGCTTTGAAATTTGCCCCAAGAGCCAGCAAGTTTGA 1140
 DB 1081 AAGTTAATGAGAGATTTCCCTCAGCTTTGAAATTTGCCCCAAGAGCCAGCAAGTTTGA 1140
 QY 1141 TTGATGATGTATGCGCAAGTTTACTTCAAGAGCGCTAGAGCAATTTACTTAAGAGTGGCT 1200
 DB 1141 TTGATGATGTATGCGCAAGTTTACTTCAAGAGCGCTAGAGCAATTTACTTAAGAGTGGCT 1200
 QY 1201 ATGAGCTCTCAGAGACAAAGTTTCCTTCTTATCTGAGTCTTTCTTCTTATATGCTGAT 1260
 DB 1201 ATGAGCTCTCAGAGACAAAGTTTCCTTCTTATCTGAGTCTTTCTTCTTATATGCTGAT 1260
 QY 1261 CGTGCTAAAGGAGAGGATTTGAGAGATTTGAGATTTAAAGAGAGCGGAGCGCTCTCT 1320
 DB 1261 CGTGCTAAAGGAGAGGATTTGAGAGATTTGAGATTTAAAGAGAGCGGAGCGCTCTCT 1320
 QY 1321 GTTCTGTTGAGAGAAAGGATGATGATGAGAGAGATTTAGTCTGCGGCTGAGAAATTTGAT 1380
 DB 1321 GTTCTGTTGAGAGAAAGGATGATGATGAGAGAGATTTAGTCTGCGGCTGAGAAATTTGAT 1380
 QY 1381 GCAAGGGGTTTCTTTTACTAGTGTCTGTTTGGTGTCTTCAAACTTTAGAGATGCA 1440
 DB 1381 GCAAGGGGTTTCTTTTACTAGTGTCTGTTTGGTGTCTTCAAACTTTAGAGATGCA 1440
 QY 1441 GATTGCTGATGATTAAGATGAGTGTGCAATGAGATTTGCGGCTGCTTGAAGGG 1500
 DB 1441 GATTGCTGATGATTAAGATGAGTGTGCAATGAGATTTGCGGCTGCTTGAAGGG 1500
 QY 1501 TCACAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
 DB 1501 TCACAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560
 QY 1561 ATTGCAAAAGGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1620
 DB 1561 ATTGCAAAAGGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1620
 QY 1621 TTGCAATTAAGCTTGAATGATTAAGTATGAGAGACATCAATATTTTATGAGTCAATGCT 1680
 DB 1621 TTGCAATTAAGCTTGAATGATTAAGTATGAGAGACATCAATATTTTATGAGTCAATGCT 1680
 QY 1681 GATTGAGCAAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
 DB 1681 GATTGAGCAAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
 QY 1741 GTTTTGTGAGAAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
 DB 1741 GTTTTGTGAGAAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
 QY 1801 GTTCTTGTGAGTTCCAATATTTGAAGTTTAAATCTATGAGTCTTGAATGATGATGATGATGAT 1860
 DB 1801 GTTCTTGTGAGTTCCAATATTTGAAGTTTAAATCTATGAGTCTTGAATGATGATGATGATGATGAT 1860
 QY 1861 TTAAGCTTAAAGCAATGCAACTTAAGAAAGTGTACTTACACAAATATTTCTATTTGCTCA 1920

Db	1861	TTAAGCTTAAGSCAATGCCAACTCTAAGAGAGGTAAGTACTACCAATATTTCTATGTCATA	1920
QY	1921	GGTATAGTTGAATCAAGTATATCAAGCGTGGAAATGATATTGAAGCTCTTGAGATGGTTTAT	1980
Db	1921	GGTATAGTTGAATCAAGTATATCAAGCGTGGAAATGATATTGAAGCTCTTGAGATGGTTTAT	1980
QY	1981	ACCTTTGGCATGGAGGATTAAGTTTTCAGCTGCTCTAGTCTCTAATCTATTCTTTAAAGATG	2040
Db	1981	ACCTTTGGCATGGAGGATTAAGTTTTCAGCTGCTCTAGTCTCTAATCTATTCTTTAAAGATG	2040
QY	2041	AGCAAGAGATCATTTGAGAGGGCAAAAACGGAAACCCAGTACCGCGTGGCATTTGTATGA	2100
Db	2041	AGCAAGAGATCATTTGAGAGGGCAAAAACGGAAACCCAGTACCGCGTGGCATTTGTATGA	2100
QY	2101	ACCCCTCCCTTGCACATTTATGTATACCTTTATGAACTCTTTATCATCATCTGAGTGCACA	2160
Db	2101	ACCCCTCCCTTGCACATTTATGTATACCTTTATGAACTCTTTATCATCATCTGAGTGCACA	2160
QY	2161	TTGATATATTTATTTCTTCACAGAAAGAGCGGCTACAAAGCAAGCTAGTGTATTATCAT	2220
Db	2161	TTGATATATTTATTTCTTCACAGAAAGAGCGGCTACAAAGCAAGCTAGTGTATTATCAT	2220
QY	2221	CAGTTATGACATGTAATGGAGACTACAAAGTTAGATCCGCGAAGAAATACCAAGATGGC	2280
Db	2221	CAGTTATGACATGTAATGGAGACTACAAAGTTAGATCCGCGAAGAAATACCAAGATGGC	2280
QY	2281	AGATCAAAAGACCAATTTGTTAGCTTGGAGAAAGACACTCTTCAGCTGCACAAAGATGAG	2340
Db	2281	AGATCAAAAGACCAATTTGTTAGCTTGGAGAAAGACACTCTTCAGCTGCACAAAGATGAG	2340
QY	2341	AAGAGAAAGCAAGATCTCTCACTTTAATGAGAGAACCCGCACTTGGCCAAAGAAATGATATA	2400
Db	2341	AAGAGAAAGCAAGATCTCTCACTTTAATGAGAGAACCCGCACTTGGCCAAAGAAATGATATA	2400
QY	2401	ACCAACAGATATAACGTCGCAAGGTTTGTACCCATGAGAAATGCACACAGTAACTCTTCAT	2460
Db	2401	ACCAACAGATATAACGTCGCAAGGTTTGTACCCATGAGAAATGCACACAGTAACTCTTCAT	2460
QY	2461	CGTATTTCTCTATCTACCGGTATGAGAGACTTTCTCTAGTCAAGAGAGACGATGACCAAGATG	2520
Db	2461	CGTATTTCTCTATCTACCGGTATGAGAGACTTTCTCTAGTCAAGAGAGACGATGACCAAGATG	2520
QY	2521	AAATATACAGCTCTTGTGTAGTATGTAATCTCGGCCCGTCAACATCTTTTCTCATCGCTCA	2580
Db	2521	AAATATACAGCTCTTGTGTAGTATGTAATCTCGGCCCGTCAACATCTTTTCTCATCGCTCA	2580
QY	2581	GAAATATCCCGGAATATATGTTGCTCACATCCACATGCGGGTTAGGAAGAAAGTGTATATG	2640
Db	2581	GAAATATCCCGGAATATATGTTGCTCACATCCACATGCGGGTTAGGAAGAAAGTGTATATG	2640
QY	2641	CATATGAACATCTGGCCCCCAATTCATATCTCTCCAGGTACAGAGCATATGACTTTCATGCAC	2700
Db	2641	CATATGAACATCTGGCCCCCAATTCATATCTCTCCAGGTACAGAGCATATGACTTTCATGCAC	2700
QY	2701	AGTACTCTCCGTCTTGGTTCACGACAGAGACATCATCACTACAGTACTCTCTCCCAATTC	2760
Db	2701	AGTACTCTCCGTCTTGGTTCACGACAGAGACATCATCACTACAGTACTCTCTCCCAATTC	2760
QY	2761	ATGAGCAACAAGTATACCATATGTTATACAAAGGGTTTACAGACATTCACACATCTGAG	2820
Db	2761	ATGAGCAACAAGTATACCATATGTTATACAAAGGGTTTACAGACATTCACACATCTGAG	2820
QY	2821	AAAGATATTTGGGTTTATCCAAATCAAAAGTCTCCCGAGATACATCATTTAGACCCCA	2880
Db	2821	AAAGATATTTGGGTTTATCCAAATCAAAAGTCTCCCGAGATACATCATTTAGACCCCA	2880
QY	2881	AATAGAGAAATGTAAATTTGTAAACAAAGCTTTTGTTTTGTCTTAAGTTAGTCAATTTAT	2940
Db	2881	AATAGAGAAATGTAAATTTGTAAACAAAGCTTTTGTTTTGTCTTAAGTTAGTCAATTTAT	2940
QY	2941	TTAACTCCCAACAGTCTCAAAATTTAATTTAATGTTTGGGCTTAAGATGCAAAATTTT	3000
Db	2941	TTAACTCCCAACAGTCTCAAAATTTAATTTAATGTTTGGGCTTAAGATGCAAAATTTT	3000

Dd		2941	TTAATCCCAACAGCTCCAAAATTAAATTAATGTTGGGGCTTAAAGAATGCCAAATTTT	3000		
Oy		3001	TTCGTCTCTGAATTCAGATTTTAGATGCTATGTATTGCTTCAGAGCTTTTAGTCACC	3060		
Dd		3001	TTTTGCTCCTGTAATTCAGATTTTAAAGATGCTAATGTATTGCTTCAGAGCTTTTAGTCACC	3060		
Oy		3061	TCAGATGCATGCATATCATCTATCAAATAAGACCTCGTGACTTGGTATCGATGCTCT	3120		
Dd		3061	TCAGATGCATGCATATCATCTATCAAATAAGACCTCGTGACTTGGTATCGATGCTCT	3120		
Oy		3121	TCATCTCTCTGCTCTGTTCCCTTCTGTTCTGCTGTCACCTGCTGAGCAATTCGCGATTCC	3180		
Dd		3121	TCATCTCTCTGCTCTGTTCCCTTCTGTTCTGCTGTCACCTGCTGAGCAATTCGCGATTCC	3180		
Oy		3181	AACCTTGCTGTACAGTTTCCCACGACAAGCTTTTCCATGATGTATTATGTCGCGCC	3240		
Dd		3181	AACCTTGCTGTACAGTTTCCCACGACAAGCTTTTCCATGATGTATTATGTCGCGCC	3240		
Oy		3241	TTCTTATCTCTGTGAGGAAGATGAATATCCCGAAGATCCAACTGTGACTTGAACATCA	3300		
Dd		3241	TTCTTATCTCTGTGAGGAAGATGAATATCCCGAAGATCCAACTGTGACTTGAACATCA	3300		
Oy		3301	TCAATTCGGAACAGAAAACAGAGCTTTTGGACATCTTTGATTTAGCATCTTTGATCTTG	3360		
Dd		3301	TCAATTCGGAACAGAAAACAGAGCTTTTGGACATCTTTGATTTAGCATCTTTGATCTTG	3360		
Oy		3361	AGGAATATCAATGAAACACTAGATACACTCACACTTGGAGGCTTTAACTGGATTTTAAAC	3420		
Dd		3361	AGGAATATCAATGAAACACTAGATACACTCACACTTGGAGGCTTTAACTGGATTTTAAAC	3420		
Oy		3421	ATGAATAGAACATTTGANTCCATGGAAATGCGTAAGACATCTACTGAGCTTCTTAAACA	3480		
Dd		3421	ATGAATAGAACATTTGANTCCATGGAAATGCGTAAGACATCTACTGAGCTTCTTAAACA	3480		
Oy		3481	AATGATATGAAGGGGTAGGGTTCATTACATGTAGTTATACAGCACTGAGATTTATGGAAG	3540		
Dd		3481	AATGATATGAAGGGGTAGGGTTCATTACATGTAGTTATACAGCACTGAGATTTATGGAAG	3540		
Oy		3541	AAAAAAGGACACAGCTTTAGATATCTACAGAGAGACAGAACATAAAGACAAAGAAAT	3600		
Dd		3541	AAAAAAGGACACAGCTTTAGATATCTACAGAGAGACAGAACATAAAGACAAAGAAAT	3600		
Oy		3601	CATTAAGTTCAGAGAGTTCGTTAAATGGCTCTATCCAATTCACATTTGGCACAAGACCAC	3660		
Dd		3601	CATTAAGTTCAGAGAGTTCGTTAAATGGCTCTATCCAATTCACATTTGGCACAAGACCAC	3660		
Oy		3661	TAAATAGATACCAAGTGGGACATGCAGAAAGAGAAATTAAGAGATATCATATCAGAGAGAG	3720		
Dd		3661	TAAATAGATACCAAGTGGGACATGCAGAAAGAGAAATTAAGAGATATCATATCAGAGAGAG	3720		
Oy		3721	AGAGATTTTTTGGAGAGGAGAGAGTTCGCCGAGGACTTCTG	3761		
Dd		3721	AGAGATTTTTTGGAGAGGAGAGAGTTCGCCGAGGACTTCTG	3761		
 RESULT 2 AAA63669 standard; cDNA: 2257 BP.						
ID	AAA63669	standard;	cDNA:	2257 BP.		
XX AC	AAA63669;					
XX XX	04-DEC-2000	(first entry)				
XX DE	cDNA sequence of the FRI gene of Arabidopsis .					
KM KM	H5L1, one locus-FRIGIDA; FRI gene; flowering time; blotting;					
KW KW	flower initiation; stem elongation; flower production; ss.					
OS OS	Arabidopsis sp.					
FT FT	Key	Location/Qualifiers				
	CDS	362..2191				
	/tag=	a				

[illegible]

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145142.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-015138.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.7%; Score 214.2; DB 21; Length 1016;
Best Local Similarity 92.6%; Pred. No. 7.5e-41; Indels 0; Gaps 0;
Matches 225; Conservative 0; Mismatches 18;

OY 2993 AAATTTTTCCTCTGATTAATTGACATTTAAGATCTAAGTATTCCTTCAGAGTTT 3052
DB 1013 AAAAATTTGCTGTGTAATTACATTTAGATGTAAGTATTCCTTCACAGTTT 954
OY 3053 AGTAACTCGATACATGCATATCATATTAATGACCTGCGCTTGGTCATCTG 3112
|||||

PR 01-JUN-1999; 99US-0137222.

[illegible]

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0143390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148561.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

of genes associated with pharmacogenomics according to one of the sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1 (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2 (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGRS (NM_003360), MRP (NM_004996, NM_019901, NM_019902, NM_019862, NM_019898, NM_019899) and their complementary sequences, or a sequence (SI) chosen from 87 sequences and their complements. The chemical pretreatment is bisulphite treatment to convert cytosines (but not methyl-cytosines) into uracils. Also included are an oligomer (II) in particular an oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in each case at least one sequence having a length of 9 nucleotides which hybridises to or is identical to a chemically pretreated DNA of genes associated with pharmacogenomics and their complements, arranged in an array for analysing diseases associated with the methylation state (CpG) and/or detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The oligomers may also be used as PCR primers. The set of 87 nucleic acids and their complements is useful for diagnosis and therapy of solid tumours and cancer. The present sequence represents one the 87 DNA sequences or its complement.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences.

Sequence 12951 BP; 3093 A; 270 C; 2979 G; 6609 T; 0 other;

Query Match 1.7%; Score 64; DB 24; Length 12951;
Best Local Similarity 50.0%; Pred. No. 0.00011;
Matches 160; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

```

QY 174 TAAACCCGCTGCTTACGACTATTTAAACATATATATACGTAATCATGCAACCT 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11261 TACACTCCAACTTAATTAACAAAAAACCCCTATCTCAATTAACAAATTTTAAAAAT 11202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 234 AACTATGTTTTCATTAATCAAAAGATTAAGAAATTAAGAAAGTGGCTAGATTCAATTAT 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11201 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 294 TTGGCATAGACTCAAAAGAGTGTATATATATCTGACTTTTATTAATTAATTAACCAAA 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11141 TTAACCTAATTAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11082
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 354 TACATATTTTCATTAAGCAAACTTAATAAGCCCTTAACATATATATGATTAACCAAGA 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11081 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 11022
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 414 AAAAGTGTTTCTCTGACTTAATAAAGATTAAGTACTTCTTAATTAATTAATTAATGT 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11021 AAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10962
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 474 GAAGTTCATATATACGATT 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10961 AATAATCACATTAACCTTTT 10942
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8

AAS63316/c
ID AAS63316 standard; DNA; 12951 BP.

AC AAS63316;

DT 29-JAN-2002 (first entry)

DE Chemically pretreated metabolism associated gene #11.

XX Human; cytostatic; anti-tumour; metabolism; metabolic disease; liver;

KW solid tumour; cancer; cytosine methylation; epigenetic; eye; kidney;

KW single nucleotide polymorphism detection; SNP; stool; urine; lung;

KW cerebral-spinal fluid; intestine; brain; prostate; breast;

KW DUSP2; EPHX2; QDPR; SGRH; SHMT2; SLC7A2; SLC7A4; TYMS; ds.

OS Homo sapiens.

PM WO200176451-A2.
XX 18-OCT-2001.
XX 06-APR-2001; 2001WO-EP04016.
XX 06-APR-2000; 2000DE-1019058.
XX 07-APR-2000; 2000DE-1019173.
XX 30-JUN-2000; 2000DE-1032529.
XX 01-SEP-2000; 2000DE-1043826.
XX (EPIC-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K.
XX WPI: 2002-010834/01.
XX New nucleic acid, useful for diagnosis and therapy of metabolic
XX disease, solid tumour and cancers, comprises segment of chemically
XX modified genomic sequences of genes associated with metabolism -
XX Claim 1; Page 45-48; 143pp; English.

The invention relates to a nucleic acid (I) comprising a sequence at least 18 bases of a segment of the chemically pretreated DNA of genes associated with metabolism such as DUSP2 (NM_004418), EPHX2 (NM_001979), QDPR (NM_000320), SGRH (NM_000199), SHMT2 (NM_005412), SLC7A2 (NM_003046), SLC7A4 (NM_004173) and TYMS (NM_001071) (all undefined). (I) are useful for diagnosis and therapy of metabolic disease, solid tumours and cancers; as primer oligonucleotides for the amplification of DNA sequences, for detecting the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) in a chemically treated DNA of genes associated with metabolism. An array of (I) is useful for ascertaining genetic and/or epigenetic parameters for the diagnosis and/or therapy of existing diseases or the predisposition to specific diseases by analysing cytosine methylations. The method involves chemically treating genomic DNA sample by a solution of bisulphite, hydrogen sulphite or disulphite such that cytosine bases which are unmethylated at the 5th-position are converted to uracil or another base which is dissimilar to cytosine in terms of hybridisation behaviour and amplifying fragments of the chemically pretreated genomic DNA. The genomic DNA is from cells or cellular components which contain DNA, sources of DNA comprising, for e.g. cell lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid, tissue embedded in paraffin such as tissue from eye, intestine, kidney, brain, heart, prostate, lung, breast or liver, histologic object slides and their combinations. Genetic parameters are mutations, in particular insertions, deletions, point mutations, inversions and polymorphisms of genes associated with metabolism and sequences further required for their regulation. Epigenetic parameters are in particular cytosine methylations and further chemical modifications of DNA bases of genes associated with metabolism. Further epigenetic parameters include for e.g. the acetylation of histones which correlates with DNA methylation. AAS63306-AAS63373 represent chemically pretreated metabolism associated genes, and related primers of the invention.

Sequence 12951 BP; 3093 A; 270 C; 2979 G; 6609 T; 0 other;

Query Match 1.7%; Score 64; DB 24; Length 12951;
Best Local Similarity 50.0%; Pred. No. 0.00011;
Matches 160; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

```

QY 174 TAAACCCGCTTGTGACTATTTAAACATATATATATGCTATATCATGCAACCT 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11261 TACACTCCAACTTAATTAACAAAAAACCCCTATCTCAATTAACAAATTTTAAAAAT 11202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 234 AACTATGTTTTCATTAATCAAAATTAAGATTAAGAAAGTGGCTAGATTCAATTAT 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11201 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 294 TTGGCATAGACTCAAAAGGCTATATATATATGCTTTTAAATTAATTAACACAA 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 11141 TTAACCTAATTAATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11082
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


OY		354	TACATATTTCCTGATTAAGAAACGTATAAACCCTTAAACGATATTAATGATTCCTCAAGA	413
Dd	11081	AA	11022	
OY		414	AAAAGTCGTTTCCTACTTAAAGATGAGTTACTTCCTCAATTAATATATATATATGCT	473
Dd	11021	AAAAATACATATATTTTATATATATACCTTCAACAACCTTATATATATATATATATC	10962	
OY		474	GAACTTCACAAATATATACAGTT	493
Dd	10961	AATATATCATATTAACCTTTT	10942	
RESULT	9			
ABK39965/c				
ID	ABK39965	standard; DNA;	15479 BP.	
XX	ABK39965;			
XX	21-MAY-2002	(first entry)		
DT		Human chemically pretreated gene sequence #23 strand 2.		
XX				
DE				
XX				
KW	Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;			
KM	Cytosatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;			
XX	UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.			
OS	Homo sapiens.			
PN	WO200202806-A2.			
PD	10-JAN-2002.			
PF	29-JUN-2001; 2001WO-EP07470.			
PR	30-JUN-2000; 2000DE-1032529.			
PR	01-SEP-2000; 2000DE-1043826.			
PA	(EPFG-) EPIDENOMICS AG.			
PI	Olek A, Piepenbrock C, Berlin K;			
DR	WPI, 2002-154757/20.			
PT	New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,			
PS	useful for detecting cytosine methylation state of genes associated			
XX	with pharmacogenomics and for therapy of diseases e.g. cancer			
CC	Claim 1; SEQ ID No 46; 24pp; English.			
CC	The invention relates to a nucleic acid comprising a sequence at			
CC	least 18 bases in length of a segment of the chemically pretreated DNA			
CC	of genes associated with pharmacogenomics according to one of the			
CC	sequences of the genes ALDH6 (NM.000693), CYP11A (NM.000781), CYP11B1			
CC	(NM.000497), CYP3A3 (NM.000776 and NM.017460), DPYD (NM.000110), EPHX2			
CC	(NM.001979), OCLN (NM.002538), TXNRD1 (NM.003330), UGT8 (NM.003360),			
CC	MRP (NM.004996, NM.019900, NM.019901, NM.019902, NM.019862, NM.019898,			
CC	NM.019899) and their complementary sequences, or a sequence (S1) chosen			
CC	from 87 sequences and their complements. The chemical pretreatment			
CC	is bisulphite treatment to convert cytosines (but not methyl-cytosines)			
CC	into uracils. Also included are an oligomer (II) in particular an			
CC	oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in			
CC	each case at least one base sequence having a length of 9 nucleotides			
CC	which hybridises to or is identical to a chemically pretreated DNA of			
CC	genes associated with pharmacogenomics and their complements, arranged in			
CC	an array for analysing diseases associated with the methylation state			
CC	(CpG) and/or detecting SNPs (single nucleotide polymorphisms)			
CC	of the 87 sequences. The oligomers may also be used as PCR primers.			
CC	The set of 87 nucleic acids and their complements is useful for diagnosis			
CC	and therapy of solid tumours and cancer. The present sequence			
CC	represents one the 87 DNA sequences or its complement.			
CC	Note: The sequence data for this patent did not form part			

Db 5630 AAAAACTATTAATATCTTTTCATTTAAATATACCAATTAATTAATAAAA 5571
 QY 370 CAAACTATTAAGCCCTTAACATATATGATTACCTCAAGAAAGTGGTTTCGC 429
 Db 5570 AAAAAACAAAAACATCAAAACATTTAAACAAAACTTTCATTAATAAAATTCATTA 5511
 QY 430 TACTTAAAGATA 442
 Db 5510 ATATTATTAATA 5498
 RESULT 12
 AAS46541/C
 ID AAS46541 standard; DNA; 10595 BP.
 XX AAS46541;
 AC AAS46541;
 XX 18-DEC-2001 (first entry)
 DE Tumour suppressor gene derived chemically modified sequence #263.
 XX Human: tumour suppressor gene; oncogene; antitumour; cytostatic;
 KM cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 XX Homo sapiens.
 OS
 XX MO200168912-A2.
 PN 20-SEP-2001.
 PD 15-MAR-2001; 2001WO-EP02955.
 PF 15-MAR-2000; 2000DE-1013847.
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX (EPIC-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 F1 WPI: 2001-602752/68.
 DR
 XX Fragments of chemically modified genes associated with tumour suppressor
 PT genes and oncogenes; useful in designing primers and probes for
 PT analysing diseases associated with cytosine methylation state e.g.
 PT cancer
 XX
 PS Claim 1; SEQ ID NO 263; 27pp; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and
 CC oncogenes having a sequence taken from 536 (actually 533 since
 CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
 CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a
 CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
 CC form part of a set of probes for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the
 CC differences serving as basis for diagnosis and/or prognosis events which
 CC are advantageous to patients. The present sequence is one of the
 CC 533 genomic sequences derived from tumour suppressor genes and
 CC oncogenes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 10595 BP; 2649 A; 176 C; 2300 G; 5470 T; 0 other;
 Query Match 1.5%; Score 56; DB 22; Length 10595;
 Best Local Similarity 49.3%; Pred. No. 0.008;
 Matches 146; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
 QY 152 TTATTTGCTGATTTGACATGTTATTAACACACTGCTTTAGGACATTTAAACATATAT 211
 Db 1274 TTTTATTTGCTGACATTAATATTCACAAAATATTTTCTCATTAATAATTTACACAT 1215
 QY 212 ACATGTCGTAATCATGCAACCTTAATCTGTTTCATTAATCAATCAAGATTAAGAG 271
 Db 1214 CCATATTAATTAACCTTAATAAAATCTAATAATCATTTATTAATTAATAAAATTCAC 1155
 QY 272 AAAAGTCGTAGATTCAATTTATTTGGCATAGACTCAAAAGAGTGATATATATCTGACTT 331
 Db 1154 AATTAATTCACACACACAGTAATTTACTTAAATAATATATATTTTCACCCCTTAA 1095
 QY 332 TTATTAATTTATTAACACAAATACATATTTTCATTAAGCAAACTATAAAGCCCTTAAC 391
 Db 1094 ATAAAAAATATATATAAAAAAATAATCTATATTAATAAAAAAATAATCTTAAAAA 1035
 QY 392 ATATATATGATTACTCAAGAAAAAGTCTTTCTCTCTACTTAAAGATAGGTTA 447
 Db 1034 AATTATATATCAACTAATAATACAAAAATTTACTCTCAATTAATAAAAAATTA 979
 RESULT 13
 ABN80061/C
 ID ABN80061 standard; DNA; 6971 BP.
 XX ABN80061;
 AC ABN80061;
 XX 15-JUL-2002 (first entry)
 DE Human chemically modified disease associated gene SEQ ID NO 78.
 XX Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
 KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
 KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
 KW antidiabetic; cytostatic; anticonvulsant; ds.
 XX Homo sapiens.
 OS Synthetic.
 XX WO200200927-A2.
 PN 03-JAN-2002.
 PD 02-JUL-2001; 2001WO-EP07536.
 PF 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX (EPIC-) EPIGENOMICS AG.
 PA Olek A, Piepenbrock C, Berlin K;
 F1 WPI: 2002-130908/17.
 DR
 XX Novel nucleic acid useful for diagnosis and therapy of diseases
 PT associated with development genes such as diabetes; comprises a
 PT sequence of a segment of chemically pretreated DNA of genes associated
 PT with development
 XX
 PS Claim 1; SEQ ID NO 78; 27pp; English.
 CC The invention relates to a nucleic acid (I) comprising a sequence at
 CC least 18 bases in length of a segment of chemically pretreated DNA (II)
 CC of genes associated with development selected from 87 genes listed in

XX	02-JUL-2001; 2001WO-EP07537.
PF	
XX	
PR	30-JUN-2000; 2000DE-1032529.
PR	01-SEP-2000; 2000DE-1043826.
XX	
PA	(EPIC-) EPIGENOMICS AG.
XX	
FI	Olek A, Plepenbrock C, Berlin K;
XX	
DR	WPI; 2002-130909/17.
XX	
PT	Nucleic acid comprising fragment of chemically modified gene, useful
PT	for diagnosis and treatment of diseases associated with abnormal
PT	cytosine methylation -
XX	
PS	Claim 1; SEQ ID NO 1571; 32pp + Sequence Listing; German.
XX	
CC	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	mucular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	rheumatoid arthritis, psoriasis and inflammatory/vulcative bowel
CC	diseases. The present sequence is a gene of the invention.
XX	
SQ	Sequence 5487 BP; 1532 A; 133 C; 1107 G; 2715 T; 0 other;
Query Match	1.5%; Score 55.4; DB 24; Length 5487;
Best Local Similarity	46.5%; Pred. No. 0.0086; Mismatches 206; Indels 0; Gaps 0;
Matches 1/9; Conservative	0; Mismatches 206; Indels 0; Gaps 0;
Y	152 TTATTGGATTGACATGTTATAAACCCTGCTTGTAGACTRTTAAACAATATATT 211
D	309 TTTTAGCTTTTAGAGTGAGTAAGCAATTAATTTTGTGTTTTTAATAATTAATAGA 368
Y	212 ACATGTCGAATCAGCACTACTATGTTTTCATTATCAATCAATCAAGATTAAGG 271
D	369 TTTATTAATTTTATTAATTTTAAATATTTTAAATTTTAAATTTATTAATATTATA 428
Y	272 AAAAGCGCAGATTCATTTATTTGGCATACCAAGAGTGTAATATATCTGACTT 331
D	429 ATTTATTAATATATATATATTTTAAATTTTAAATTTTAAATTTTAAATTTAATA 488
Y	332 TTATTAATTTATTAACAACATACATTTTTCATAAGCAAACATAAAGCCCTAAC 391
D	489 TTAATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAGAT 548
Y	392 ATTATATGATTAACCCAAGAAAAGCGTTTCTCCTACTTAATAAAGATAGTACTGC 451
D	549 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTATA 608
Y	452 CTATTAATTAATTAATTTATTTGTAACCTCACATATACAGTTCATTAATTTGGTAAT 511
D	609 AATATGATTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTTTGTGTTTA 668
Y	512 TGACCGATTTAAGGAGCTGGCAAT 536
D	669 GGATTTTGTGTTTGAAGATTAAT 693
RESULT 15	
ID	AAS46571/c
XX	AAS46571 standard; DNA; 8770 BP.
AC	AAS46571;
DT	18-DEC-2001 (first entry)
XX	
DE	Tumour suppressor gene derived chemically modified sequence #293.
XX	
KW	Human; tumour suppressor gene; oncogene; antitumour; cytostatic;

KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 OS Homo sapiens.
 PN WO200168912-A2.
 XX
 XX
 PD 20-SEP-2001.
 XX
 XX
 PF 15-MAR-2001; 2001WO-EP02955.
 XX
 XX
 PR 15-MAR-2000; 2000DE-1013847.
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 XX
 DR WPI; 2001-602752/58.
 XX
 XX
 PT Fragments of chemically modified genes associated with tumour suppressor
 PT genes and oncogenes, useful in designing primers and probes for
 PT analysing diseases associated with cytosine methylation state e.g.
 PT cancer
 XX
 XX
 PS Claim 1; SEQ ID No 293; 27pp; English.
 CC The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and
 CC oncogenes having a sequence taken from 536 (actually 533 since
 CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
 CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a
 CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
 CC form part of a set of probes for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the
 CC differences serving as basis for diagnosis and/or prognosis events which
 CC are disadvantageous to patients. The present sequence is one of the
 CC 533 genomic sequences derived from tumour suppressor genes and
 CC oncogenes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.
 XX
 XX
 SQ Sequence 8770 BP; 2335 A; 213 C; 2112 G; 4110 T; 0 other;
 Query Match 1.58; Score 55.4; DB 22; Length 8770;
 Best Local Similarity 49.28; Pred. No. 0.01;
 Matches 146; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
 QY 205 ATATTTACATGCTGTATATCATGACCTACATGTTTCATTAATCAATCAAGAA 264
 DB 2352 ATATTAACATTAACATTAACCTTATCAAAATTAACAAACCTTTAAAAAATA 2473
 QY 265 TAAAGAGAAAGTCGTAGATTCAATTATTTGGCATAGACTCAAAAGAGTATATAT 324
 DB 2472 AAAAAAAAAAATAATTAATAATACCTCTATATAAAAAAAAAAATACATTTA 2413
 QY 325 CTGACTTTTAAATTTTAAACCAATATTTTCTTACAGCAAACTATTAAGC 384
 DB 2412 AAAAAATTTATTAATTAATTAACCAAAATTAATAATTAATAAATCTTAATA 2353
 QY 385 CCTAAACATTAATGATTAACCTCAAGGAAAAAGTCGTTTCTCCTACTTAAGATAG 444

DB 2352 AATTTAATTTAACTATAAACCTTATCAAAATTAACAAACCTTTAAAAAATA 2293
 QY 445 TTACTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 501
 DB 2292 TTTATTAACATTTAATAATTAACCTTAATAAAAAAAAAATTAATTAATTAATA 2236

Search completed: July 11, 2003, 10:41:51
 Job time : 762 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2003, 10:28:59 ; Search time 151 Seconds
(without alignments)
7638.485 Million cell updates/sec

Title: US-09-890-475-2

Perfect score: 3761

Sequence: 1 agtactacaagtcacact.....aagtcgcgagagctctg 3761

Scoring table: Gapped 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/PCITUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	1.7	7218	1 US-08-232-463-14	Sequence 14, Appl
2	56.8	1.5	7218	1 US-08-232-463-14	Sequence 14, Appl
3	48.6	1.3	19124	2 US-08-487-8268-13	Sequence 13, Appl
4	46.8	1.2	8920	2 US-08-446-855A-1	Sequence 1, Appl
5	46.8	1.2	8920	4 US-09-150-741-1	Sequence 1, Appl
6	44.4	1.2	168575	4 US-09-426-290-1	Sequence 1, Appl
7	44.4	1.2	9048	3 US-08-973-273-4	Sequence 4, Appl
8	43.8	1.2	837	4 US-08-998-416-288	Sequence 288, App
9	43	1.1	636	4 US-08-998-416-1137	Sequence 1137, Ap
10	42.4	1.1	1939	1 US-07-715-731B-2	Sequence 2, Appl
11	41.6	1.1	8920	2 US-08-446-855A-1	Sequence 1, Appl
12	41.6	1.1	8920	4 US-09-150-741-1	Sequence 1, Appl
13	41.4	1.1	6243	4 US-09-056-075-1	Sequence 1, Appl
14	40.8	1.1	6124	4 US-08-213-419B-3	Sequence 3, Appl
15	40.6	1.1	4673	1 US-07-638-431-1	Sequence 1, Appl
16	40.6	1.1	4673	1 US-07-638-431-1	Sequence 1, Appl
17	40.6	1.1	20674	4 US-09-641-638-651	Sequence 651, App
18	40.4	1.1	636	4 US-08-998-416-1137	Sequence 1137, Ap
19	40.4	1.1	2110	4 US-09-419-459-1	Sequence 1, Appl
20	40.4	1.1	6152	4 US-08-973-462-1	Sequence 1, Appl
21	40.2	1.1	615	4 US-08-998-416-186	Sequence 186, App
22	40.2	1.1	658	4 US-08-998-416-595	Sequence 595, App
23	40.2	1.1	3095	6 5231168-1	Patent No. 5231168
24	40.2	1.1	6243	2 US-09-056-075-1	Sequence 1, Appl
25	40.2	1.1	19124	2 US-08-487-8268-13	Sequence 13, Appl
26	40	1.1	3925	2 US-09-047-026A-3	Sequence 3, Appl
27	39.8	1.1	665	2 US-08-883-795A-36	Sequence 36, Appl

28	39.8	1.1	5852	1 US-07-867-106-2	Sequence 2, Appl
29	39.6	1.1	665	2 US-08-883-795A-36	Sequence 36, Appl
30	39.6	1.1	3167	4 US-09-276-531-132	Sequence 132, App
31	39.6	1.1	12730	4 US-09-004-838-91	Sequence 91, Appl
32	39.4	1.0	6124	4 US-08-213-419B-3	Sequence 3, Appl
33	39.2	1.0	1983	4 US-09-453-702B-36	Sequence 36, Appl
34	39	1.0	950	4 US-09-593-995-3	Sequence 3, Appl
35	38.8	1.0	1391	2 US-08-950-168-2	Sequence 2, Appl
36	38.8	1.0	1391	4 US-09-365-705-2	Sequence 2, Appl
37	38.8	1.0	1404	4 US-09-257-179-34	Sequence 34, Appl
38	38.8	1.0	2287	4 US-08-845-258-8	Sequence 8, Appl
39	38.8	1.0	2287	4 US-08-990-571-8	Sequence 8, Appl
40	38.8	1.0	2287	4 US-08-723-142A-8	Sequence 8, Appl
41	38.8	1.0	2287	4 US-09-528-784A-8	Sequence 8, Appl
42	38.8	1.0	6921	4 US-09-643-587-117	Sequence 117, App
43	38.6	1.0	2429	4 US-09-386-493-3	Sequence 3, Appl
44	38.4	1.0	319	1 US-07-593-657-14	Sequence 14, Appl
45	38.4	1.0	837	4 US-08-998-416-288	Sequence 288, App

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 304772/114 INMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-F1s
US-08-232-463-14
Query Match 1.7%, Score 63; DB 1; Length 7218;

ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/467,826B
FILING DATE: 10 SEP 1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned 29, 655
REGISTRATION NUMBER: N1H121.001CP1
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
AMT-SENSE: NO
US-08-487-826B-13

Query Match 1.3%; Score 48.6; DB 2; Length 19124;
Best Local Similarity 46.2%; Pred. No. 0.014;
Matches 162; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

QY 172 TATAACACGCTGCTGATGATTTAAACATATATTACATGCTGATATGATGAC 231
DB 15621 TAAAAAATGATTTATATATATATATATATATATATATATATATATAT 15680
QY 232 CTACTATGTTTATATATATATATATATATATATATATATATATATATAT 291
DB 15661 AAAAAAAT 15740
QY 292 ATTTGGCATAGCTCAAAAGAGTATATATATATATATATATATATATATAT 351
DB 15741 AT 15800
QY 352 AATCATATTTTCTATAGCAAACTATATATATATATATATATATATATATAT 411
DB 15801 AAAAAAAT 15860
QY 412 GAAAAAGTGTTCCTTCTACTTAAAGATAGTTTCTTCAATTAATATATATAT 471
DB 15861 AAT 15920
QY 472 GGAAGCTCACAATATATAGTCAATATATATATATATATATATATATATAT 522
DB 15921 AAAAAAAT 15971

RESULT 4

US-08-446-855A-1/c
Sequence 1, Application US/08446855A
Patent No. 5849573
GENERAL INFORMATION:
APPLICANT: Stewart, Thomas S
APPLICANT: Flores, Maria V
APPLICANT: O'Sullivan, William J
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
NUMBER OF SEQUENCES: 2
TITLE OF INVENTION: phosphate synthetase II
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 1100 No. 5849573th Glabe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 47-80
TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-816-4000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic
US-08-446-855A-1

Query Match 1.2%; Score 46.8; DB 2; Length 8920;
Best Local Similarity 45.5%; Pred. No. 0.03; Mismatches 247; Indels 5; Gaps 1;
Matches 210; Conservative 0; Mismatches 247; Indels 5; Gaps 1;

QY 105 TAGATAGACAGTACTTGTGGCTCATATATACGACGACGACGACGATATATGATTA 164
DB 8877 TACAT 8818
QY 165 GACATGTTATTAACCACTGCTTACGACATATATATATATATATATATATATAT 224
DB 8817 AAAAAATCAT 8758
QY 225 ATGCACCTACATATGTTTCTATATATATATATATATATATATATATATATATAT 284
DB 8757 AT 8698
QY 285 TTCAATTA-----TTGGCATAGCTCAAAAGAGTATATATATATATATATATAT 339
DB 8697 TTGAAAAACCATTTGGTTTATATATATATATATATATATATATATATATATAT 8638
QY 340 TTATTAACCAAT 399
DB 8637 AAAAAATTAAT 8578
QY 400 ATTACTCAAGAGGAAAAAGTGTTCCTTCTACTTAAAGATAGTTTCTTCTAATTA 459
DB 8577 TAATATATTAACGTAATATATATATATATATATATATATATATATATATATAT 8518
QY 460 TAT 519
DB 8517 AT 519
QY 520 TTAAGGAGCTGGAATATAGGCTTCTGCAATCTTTTTCCTT 561
DB 8457 GTAT 8416

RESULT 5

US-09-150-741-1/c
Sequence 1, Application US/09150741
Patent No. 6183996
GENERAL INFORMATION:
APPLICANT: Stewart et al.
TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl phosphate
Patent No. 6183996
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/150,741
EARLIER FILING DATE: 1998-09-10
EARLIER APPLICATION NUMBER: PL6380
EARLIER FILING DATE: 1992-12-16
EARLIER APPLICATION NUMBER: A093/00617
EARLIER FILING DATE: 1993-12-02
EARLIER APPLICATION NUMBER: 08/446,855
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 8920
TYPE: DNA
ORGANISM: Plasmodium falciparum
US-09-150-741-1

BEST Local Similarity 49.6%; Pred.NO. 0.5;
Matches 140; Conservative 0; Mismatches 141; Indels 1; Gaps 1;

OY 224 CAGCACCCTACTACTGTGTTTTCATTATACCAATATCCAAATATAAGAAGAAAATCGCTAG 283
Db 108924 CATATATATTAATATATATATATATGAATGTTATATATATATATATATATATATATGTA 108933

OY 284 ATTCATTTTTGGCATGACCTCAAAAGAGTGTAATATATCAGCTTTTATPMAAATAT 343
Db 108984 TTATACTTAT 109043

OY 344 TAAACCAAATACATATTTTTCATATAGCAAAACTATAAAAAGCCCATAAACATATATGCATTA 403
Db 109044 TACATATTAATATCATATATATAT -TATATTTTATATATATATATATATATATATATATA 109102

OY 404 CCCAAGAGAAAAGTCGTTTCCTCCACTTAAAGATAGTGTAATCTCTCTTAATTAATATA 463
Db 109103 ACCTA 109162

OY 464 TAATTTATGTGAACCTCACCAATATACACTTCATATATTTTG 505
Db 109163 TAACCTATG 109204

RESULT 7
US-08-973-273-4
Sequence 4: Application US/08973273
Patent No. 6140085
GENERAL INFORMATION:
APPLICANT: Dean, Caroline
APPLICANT: MacKnight, Richard C
APPLICANT: Bancroft, Ian
APPLICANT: Lister, Clare K
TITLE OF INVENTION: Genetic Control of Flowering
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vandenberg P.C.
STREET: 1100 NO. 6140085th Glebe Road, 8th Floor
City: Arlington
STATE: Virginia
COUNTRY: United States of America
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973.273
FILING DATE: 01-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/01332
FILING DATE: 03-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9511196.9
FILING DATE: 02-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ms Mary J Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-29
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 9048 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPERTEXTAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:

ORGANISM: Brassica
US-08-973-273-4

Query Match 1.2%; Score 44; DB 3; Length 9048;
Best Local Similarity 45.4%; Pred. No. 0.16;
Matches 158; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

194 TATTAAACAAATATATACATGCTGATCATGCAACCACTGATGTTTCATTATCA 253
130 TTTTCAGATAGTAAATCAACATCAAGATATATATATGTCACAAATTTTATCAAAA 189
254 AATCAAGAAATTAAGAGAAAAGTCGATGATTCATATTTGCGATAGACTCAAAAG 313
190 TATATACCTATATATGTTTAAATTTTAAACACATATCTAGCAAAATGTTAGAA 249
314 TGTATATATCTGCTTTTATATATATTAATTAACACAAATATTTTCATAGCAAA 373
250 AATATCTTATATATATGTTTAACTTTTATATATTAATAAATTTGTTTACATAGCA 309
374 ACTATAAAGCCCTTAACATATATATGATTTACCTCAAGAAAAGTCGTTTCTCTACT 433
310 CAATATATATATAGATTAATTAATTTTAAATTTTGAATATTTATATATATTTA 369
434 TAAAGATAGTACTTCTCTATATATATATATTTATGTCAGACTTCACATATAGCT 493
370 AGAATCAAT 429
494 CAATAAATTTGTTATTTGACCGATTTAAGAGAGATGGAATTTAGG 541
430 TTAAT 477

RESULT 8
US-08-998-416-288/c
Sequence 288, Application US/08998416
Patent No. 6239264

GENERAL INFORMATION:
APPLICANT: Philippesen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYPII
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: CH 0016/97
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Weiss, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 288:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1241RP
US-08-998-416-288

Query Match 1.2%; Score 43.8; DB 4; Length 837;
Best Local Similarity 48.0%; Pred. No. 0.06;
Matches 157; Conservative 0; Mismatches 167; Indels 1; Gaps 1;

171 TTTAAACCACTGCTTTAGTACTATTTAAACATATATATGATGCTATATCA 230
589 TTAAT 530
231 CTAATCTGTTTCTTATATCAATACAA-CAATTAAGAGAAAGTCGATGCTCA 289
529 ATGAAT 470
290 TTAATGCGATGACTCAAAAGAGTATATATATCTGACTTTTATATATATTAACA 349
469 ATTTATATATATATATCTTTTAAAGATTAATATATATATATATATATATATA 410
350 CAATATATATTTTCAATACCAAACTATATTAAGCCCTTAACATATATATATCTCA 409
409 AATGAT 350
410 AGGAAAGTCGTTTCTCTCTACTTAAAGATGATGTTACTCTTATATATATATAT 469
349 AATTAAT 290
470 ATGTGAATCTCACATATATACAT 492
289 TTTAAT 267

RESULT 9
US-08-998-416-1137
Sequence 1137, Application US/08998416
Patent No. 6239264

GENERAL INFORMATION:
APPLICANT: Philippesen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYPII
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: CH 0016/97

FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meligs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1137:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1692RP
US-08-998-416-1137

Query Match 1.1%; Score 43; DB 4; Length 636;
Best Local Similarity 44.9%; Pred. No. 0.086;
Matches 163; Conservative 0; Mismatches 200; Indels 0; Gaps 0;

QY 147 AACCGTATTGATGACAGCTGTTAAACCACTGCTTAGGACTATTAAACAAT 206
DB 247 AATATATTAGATATTATTATTCTTTAAATTAATTAATTAATTAATTAATTA 306
QY 207 ATATACATGTCGATACGACACCTACATGCTTTCAATTAATTAATTAATTAATTA 266
DB 307 ATATATATTATTATATGTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 366
QY 267 AAGAGAAAGTGCCTAGATTCATTTATTTGGCATTAGCTCAAAAGAGTATATATCT 326
DB 367 TATTTAAATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGTT 426
QY 327 GACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 386
DB 427 GATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 486
QY 387 TAAACATATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATGTT 446
DB 487 TCTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 546
QY 447 ACTTCCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 506
DB 547 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 606
QY 507 TAA 509
DB 607 TAA 609

RESULT 10
US-07-715-751B-2
Sequence 2, Application US/0715751B
Patent No. 5391725
GENERAL INFORMATION:
APPLICANT: CORUZZI, GLORIA M
APPLICANT: EDWARDS, JANICE W
APPLICANT: WALKER, ELSETH L
APPLICANT: BREARS, TIMOTHY B
TITLE OF INVENTION: NOVEL ORGN-SPECIFIC PLANT PROMOTER
TITLE OF INVENTION: SEQUENCES
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: N.Y.
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/715,751B
FILING DATE: 19910613
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MISROCK, S. LESLIE
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 3288-017-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1939 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-715-751B-2

Query Match 1.1%; Score 42.4; DB 1; Length 1939;
Best Local Similarity 49.6%; Pred. No. 0.21;
Matches 138; Conservative 0; Mismatches 136; Indels 4; Gaps 1;

QY 225 ATGCACCTATGATGCTTTCATTAATCAAAATCAAAAGATTAAGAGAAAGGCGGAGA 284
DB 1209 ATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1268
QY 285 TTCAATATTTGGCATGACTCAAAAGAGTATTAATTAATTAATTAATTAATTAATTAAT 344
DB 1269 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1324
QY 345 AAACCAATATGATTTTCATTAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTA 404
DB 1325 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1384
QY 405 CTCAGAGAAAGTGCCTTTTCCTTCACTTAAGATAGGTTCTCTTAATTAATTAAT 464
DB 1385 ATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1444
QY 465 AATTATGTAACCTTCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 502
DB 1445 GTATCTACCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1482

RESULT 11
US-08-446-855A-1
Sequence 1, Application US/08446855A
Patent No. 5849573
GENERAL INFORMATION:
APPLICANT: Stewart, Thomas S
APPLICANT: Flores, Maria V
APPLICANT: O'Sullivan, William J
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
TITLE OF INVENTION: phosphate synthetase II
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 1100 No. 5849573th Glebe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:


```

; LOCATION: 3770..4013
; OTHER INFORMATION: /note= "Rp4 origin of DNA transfer (oriT) from
; OTHER INFORMATION: plasmid Rp4"
US-09-056-075-1

```

Query Match	1.18;	Score 41.4;	DB 2;	Length 6243;
Best Local Similarity	46.68;	Pred. No. 0.66;		
Matches 132;	Conservative	0;	Mismatches 151;	Indels 0;
			Gaps	0.

QY	229	AACTCACTAATGTTTCCTTATCATAAATCAACAGATTAAGAGAAAGGTGGGTGATYCA	288
Db	1172	AGCCTAAAAAATGAGGGTAAAAATAAAAAATAAAAAATAAAAAATAAAAAATPAA	1231
QY	289	ATTATTGGCCATAGACTCAAAAGAGTGATATATCTGACTTTTATTAAATTATTAAC	348
Db	1232	AAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA	1291
QY	349	ACAATACATATTTTCATACAGCAAACTATAAAAAGCCCTAAACATATATGTTACTCA	408
Db	1292	ATAAAAAATTTAAAAATAAAAAATATAAAAATAAAAAATATAAAAAATAAAAAAT	1351
QY	409	AAGGAAAAAGTCGTTTCTCTACTTAAAGATAGTTACTTCTTAATTAATATTAATT	468
Db	1352	AAAAATAAAAAATAAAAAATATPAAATATAAAAAATAAAAAATAAAAAATATATTTT	1411
QY	469	TATGGAAGCTCACAAATATACAGTCACATCAAAATTTGGTAAT	511
Db	1412	TATTTAAAGTTGAAAAAATTTTATTTATATATATAATCTTT	1454

```

RESULT 14
US-08-213-419B-3
; Sequence 3, Application US/08213419B
; Patent No. 6333406
; GENERAL INFORMATION:
; APPLICANT: Ineslbury, J. et al.
; TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: J11-002CNCPC
; CURRENT APPLICATION NUMBER: US/08/213,419B
; CURRENT FILING DATE: 1994-03-14
; PRIOR APPLICATION NUMBER: US 07/870,506
; PRIOR FILING DATE: 1992-04-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 6124
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2407)..(2439)
; NAME/KEY: CDS
; LOCATION: (2598)..(3404)
; NAME/KEY: CDS
; LOCATION: (3580)..(3720)
; NAME/KEY: CDS
; LOCATION: (3850)..(5835)
US-08-213-419B-3

```

[illegible]

Db 2179 ATTTAAATTAATTAATTAATTTATTAATTAAAT-TTATATATATATATATATATATTAATTTT 2237
 QY 406 TCAAAAGAAAAAGTCGTTTCTCTACTTAAAGATAGTTACTTCCATTAATATATA 465
 Db 2238 TTAATTTAAATTAATTAATAGTGTCTCCAAAAAATAAAAAATAAATATATATATA 2297
 QY 466 ATTTATGTGAACTTCACAAATATACAGTTCATTAATAATTTGTAATTTGACCGATTTAAG 525
 Db 2298 TATATATAAATACATATATTTTAATACATTAACGAAAAATTAATTAATCAAAACATATTCAA 2357
 QY 526 AGAGTGGAAATT 537
 Db 2358 AAAATTAAGTT 2369

RESULT 15
 US-07-638-431-1/c
 Sequence 1, Application US/07638431
 Patent No. 5198535
 GENERAL INFORMATION:
 APPLICANT: Hoffman, Stephen L.
 APPLICANT: Charoenvit, Yupin
 APPLICANT: Hustedrom, Richard
 APPLICANT: Khumsmith, Srisin
 APPLICANT: Rogers IV, William O.
 TITLE OF INVENTION: Protective malaria sporozoite surface protein
 TITLE OF INVENTION: Immunogen and gene
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: A. David Spevack
 STREET: NMRDC Building 1 T-12 National Naval
 STREET: Medical Center
 CITY: Bethesda
 STATE: MD
 COUNTRY: USA
 ZIP: 20814-5044
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.24
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/638,431
 FILING DATE: 19910110
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Spevack, Avrom D.
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 295-6759
 TELEFAX: (301) 295-4033
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4673 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: N
 ANTI-SENSE: N
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium yoelii
 STRAIN: 17X(NL)
 DEVELOPMENTAL STAGE: erythrocytic stage
 TISSUE TYPE: Blood
 CELL TYPE: erythrocytic stage
 IMMEDIATE SOURCE:
 LIBRARY: Py-lambdaagtl1-2-7 kb genomic expression
 CLONE: Py10.1111
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 718..3195
 OTHER INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2003, 10:44:33 ; Search time 638 Seconds

(without alignments)
9308.153 Million cell updates/sec

Title: US-09-890-475-2

Perfect score: 3761
Sequence: 1 agtactcaagtcacact.....aagtcgcgaggtctctg 3761

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1105431 seqs, 789497651 residues

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PC1_NEM_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCRTS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95.4	2.5	356	10	US-09-770-791-607
2	94.4	2.5	271	10	US-09-924-035A-6
3	54.4	1.4	5689	9	US-10-239-676-90
4	53.4	1.4	5689	9	US-10-239-676-89
5	53.4	1.4	11036	9	US-10-239-676-117
6	52.8	1.4	3991	9	US-10-074-045-60
7	52	1.4	8085	9	US-10-172-086-39
8	51.8	1.4	6306	9	US-10-239-676-224
9	51.2	1.4	3001	9	US-10-172-086-58
10	50.8	1.4	2000	9	US-09-938-842A-3198
11	50.8	1.4	2004	10	US-09-887-576-264
12	50	1.3	960	9	US-10-198-846-6381
13	49.2	1.3	6397	9	US-10-239-676-107
14	48.8	1.3	53332	9	US-10-224-562-3
15	48.8	1.3	53332	10	US-09-801-861-3
16	48.6	1.3	6246	9	US-10-172-086-27
17	48.4	1.3	14708	9	US-10-239-676-221
18	48.2	1.3	15732	9	US-10-239-676-95
19	48	1.3	566	10	US-09-828-644-42

20	48	1.3	640681	10	US-09-790-988-1	Sequence 1, Appl
21	47.6	1.3	3991	9	US-10-074-045-60	Sequence 60, Appl
22	47.4	1.3	393	10	US-09-960-352-4582	Sequence 4582, Ap
23	47.2	1.3	6665	9	US-10-239-676-4	Sequence 4, Appl1
24	47.2	1.3	146547	9	US-10-017-128-1	Sequence 1, Appl1
25	47	1.2	7934	9	US-10-239-676-154	Sequence 154, Appl
26	46.8	1.2	413	10	US-09-960-352-2919	Sequence 2919, Ap
27	46.8	1.2	7823	9	US-10-239-676-198	Sequence 198, App
28	46.6	1.2	6620	9	US-10-239-676-196	Sequence 196, App
29	46.6	1.2	17419	9	US-10-239-676-100	Sequence 100, App
30	46.4	1.2	2265	9	US-09-764-872-939	Sequence 939, App
31	46.2	1.2	4632	9	US-10-024-623-7	Sequence 7, Appl1
32	46.2	1.2	302250	10	US-09-962-832-154	Sequence 154, App
33	46	1.2	2000	9	US-09-938-842A-4722	Sequence 4722, Ap
34	46	1.2	5001	9	US-10-172-086-70	Sequence 70, Appl
35	46	1.2	7001	9	US-10-172-086-59	Sequence 59, Appl
36	46	1.2	11047	9	US-10-239-676-188	Sequence 188, App
37	45.8	1.2	5659	9	US-10-172-086-31	Sequence 31, Appl
38	45.6	1.2	2000	9	US-09-938-842A-4610	Sequence 4610, Ap
39	45.6	1.2	9293	9	US-10-239-676-106	Sequence 106, App
40	45.6	1.2	15832	9	US-10-224-562-3	Sequence 3, Appl1
41	45.6	1.2	53332	9	US-09-801-861-3	Sequence 3, Appl1
42	45.6	1.2	53332	10	US-09-801-861-3	Sequence 3, Appl1
43	45.4	1.2	6306	9	US-10-239-676-129	Sequence 129, App
44	45.4	1.2	7657	9	US-10-239-676-186	Sequence 186, App
45	45.4	1.2	640681	10	US-09-790-988-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-770-791-607/c
Sequence 607, Application US/09770791
Patent No. US2002062014A1
GENERAL INFORMATION:
APPLICANT: Gorlach, Jörn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Matthew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Woessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krickler, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2029 (PARA-018PRV)
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,480
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: PstSeq for Windows Version 4.0
SEQ ID NO 607
LENGTH: 356
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-791-607
Query Match 2.5%: Score 95.4: DB 10: Length 356:
Best Local Similarity 99.0%: Pred. No. 6.5e-13:
Matches 96: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

QY	3139	CCCTCTGTTGTCGTTGACACGTCGACACATTTGGATTTCAACCTTGTTGTTACAGTT	3198
Db	356	CCCTCTGTTTCTCTTGACACGCTCGACCAATTTGGGATTTCAACCTTGTTTACAGTT	297
QY	3199	TCCCATGACACAAAGCTTTTCATGAAATGATTTAGCT	3235
Db	296	TCCCATGACACAAAGCTTTTCATGAAATGATTTATGCT	260

```

1 RESULT 2
2 US-09-924-035A-6/c
3 : Sequence 6, Application US/09924035A
4 : Patent No. US20020142319A1
5 :
6 : GENERAL INFORMATION:
7 :
8 : APPLICANT: Grilach, Jrm
9 :
10 : TITLE OF INVENTION: Expressed Sequences of Arabidopsis
11 :
12 : TITLE OF INVENTION: thaliana
13 :
14 : FILE REFERENCE: 2011US
15 :
16 : CURRENT APPLICATION NUMBER: US/09/924,035A
17 :
18 : CURRENT FILING DATE: 2000-08-11
19 :
20 : PRIOR APPLICATION NUMBER: US 60/148,784
21 :
22 : PRIOR FILING DATE: 1999-08-13
23 :
24 : NUMBER OF SEQ ID NOS: 900
25 :
26 : SOFTWARE: FastSeq for Windows Version 3.0
27 :
28 : SEQ ID NO 6
29 :
30 : LENGTH: 271
31 :
32 : TYPE: DNA
33 :
34 : ORGANISM: Arabidopsis thaliana
35 :
36 : US-09-924-035A-6

```

```

Query Match Similarity      2.5%; Score 94.4; DB 10; Length 271;
Best Local Similarity      99.0%; Pred. No. 9, 5e-13;
Matches      95; Conservative      0; Mismatches      1; Indels      0; Gaps      0

QY      3140 CTTCTGTTCTCGTTGCACTGCTCGAGCAATTGCGGANTCCAACTTGTCCTTACAGTTT 3199
|||||
DB      271 CTTCTGTTCTCTTTTGCACTGCTCGAGCAATTGCGGANTCCAACTTGTCCTTACAGTTT 212
|||||

QY      3200 CCCATGACACAGCTTTTCCATGATGATTTATGTT 3235
|||||
DB      211 CCCATGACACAGCTTTTCCATGATGATTTATGTT 176
|||||

RESULT 3
US-10-239-676-90/c
; Sequence 90, Application US/10239676
; Publication No. US20030082609A1
;
GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239, 676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 90
; LENGTH: 5689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

```

```

?      FEATURE:
?      NAME/KEY: unsure
?      LOCATION: (1818, 1930..1931, 1934..1935, 1947..1948, 1966)
?      US-10-239-676-90

```

[illegible]

```

RESULT 4
US-10-239-676-89/c
; Sequence 89, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 89
; LENGTH: 5689
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3724, 3742..3743, 3755..3756, 3759..3760, 3872)
US-10-239-676-89

Query Match          1.4%; Score 53.4; DB 9; Length 5689;
Best Local Similarity 47.7%; Pred No. 0.048;
Matches 156; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

OY      161 ATTAGCAGTGTATTAATTAACCCACGCTTGTAGTACTATTAAAACAATATATTACATGTCGT 220
        I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
Db      1568 ACTATATACCTCTCTAAAAATAAATTTTATATATATATCACTATCTATCTCAATAAAACT 1508
        I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
221 AATCATGCACCTACTACTAGTTTCATTTATCAATATACAAATAAAGAGAAAAGTGGC 280

```



```

; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3198
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-3198

Query Match      1.4%; Score 50.8; DB 9; Length 2000;
Best Local Similarity 48.3%; Pred. No. 0.11; Mismatches 152; Indels 0; Gaps 0;
Matches 142; Conservative 0;

QY      200  AACAAATATATTCATGTCGTATATCATGCAACCTACTAGTTTCTATTAATCAATACA 259
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db      266  AATATATTTATCTCTTGTGAGATTAAGTCAAAATCAATTAATTAATTAATTAATTA 325

QY      260  AAGATTAAGAGAAAAGTGGCTAGATTCATTAATTTGGCATAGACTCAAAAGAGTGATA 319
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db      326  ATTAAAAACGAATATCAATTAATTAATTTTCAAAAATATATACCAATCAAAATCAACATA 385

QY      320  TATATCGACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 379
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db      386  AAGATTGATATCTAAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 445

QY      380  AAAGCCCTAAACATATATATGATTAATCTCAAGAAAAGAGTGTCTCTCACTTAATAAG 439
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db      446  AATTCCTCAAAAAGTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 505

QY      440  ATAGGTTACTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 493
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db      506  ATTATTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 559

RESULT 11
US-09-887-576-264
; Sequence 264, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887, 576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213, 848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214, 087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258, 692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 264
```

```

; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-887-576-264

Query Match      1.4%; Score 50.8; DB 10; Length 2004;
Best Local Similarity 48.3%; Pred. No. 0.11; Mismatches 152; Indels 0; Gaps 0;
Matches 142; Conservative 0;

QY      200  AACAAATATATTCATGTCGTATATCATGCAACCTACTAGTTTCTATTAATCAATACA 259
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db      266  AATATATTTATCTCTTGTGAGATTAAGTCAAAATCAATTAATTAATTAATTAATTAATTA 325

QY      260  AAGATTAAGAGAAAAGTGGCTAGATTCATTAATTTGGCATAGACTCAAAAGAGTGATA 319
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db      326  ATTAAAAACGAATATCAATTAATTAATTTTCAAAAATATATACCAATCAAAATCAACATA 385

QY      320  TATATCGACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 379
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db      386  AAGATTGATATCTAAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 445

QY      380  AAAGCCCTAAACATATATGATTAATCTCAAGAAAAGAGTGTCTCTCACTTAATAAG 439
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db      446  AATTCCTCAAAAAGTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 505

QY      440  ATAGGTTACTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 493
      111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db      506  ATTATTTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 559

RESULT 12
US-10-198-846-6381/C
; Sequence 6381, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinhann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198, 846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306, 220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6381
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 8, 66, 71, 79, 81, 84, 85, 86, 88, 117, 119, 125, 134,
; LOCATION: 149, 156, 162, 166, 172, 179, 180, 194, 230, 249, 250, 253,
; LOCATION: 268, 271, 273, 274, 275, 276, 277, 286, 291, 299, 312, 333,
; LOCATION: 359, 366, 374, 393, 404, 406, 409, 411, 422, 424, 427
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 428, 433, 450, 455, 466, 468, 471, 487, 510, 512, 513, 516,
; LOCATION: 517, 518, 528, 530, 534, 539, 553, 555, 559, 575, 580, 588,
; LOCATION: 590, 610, 625, 628, 642, 643, 656, 658, 659, 675, 685, 686,
; LOCATION: 690, 700, 702, 706, 716, 731, 745, 748, 756, 798, 802
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 807, 808, 819, 820, 830, 838, 840, 844, 845, 849, 861,
; LOCATION: 871, 874, 875, 891, 892, 893, 898, 900, 901, 902, 904, 907,
; LOCATION: 917, 923, 925, 926, 927, 929, 939
; OTHER INFORMATION: n = A,T,C or G
```


